ROUTING ALGORITHMS FOR LARGE SCALE WIRELESS SENSOR NETWORKS

A Thesis

by

LAKSHMANA PRASANTH NITTALA VENKATA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2004

Major Subject: Computer Science

ROUTING ALGORITHMS FOR LARGE SCALE WIRELESS SENSOR NETWORKS

A Thesis

by

LAKSHMANA PRASANTH NITTALA VENKATA

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Dmitri Loguinov (Chair of Committee)	Andreas Klappenecker (Member)
Madhav Pappu (Member)	Valerie E. Taylor (Head of Department)

December 2004

Major Subject: Computer Science

ABSTRACT

Routing Algorithms for Large Scale Wireless Sensor Networks. (December 2004)

Lakshmana Prasanth Nittala Venkata, B.Tech, Andhra University

Chair of Advisory Committee: Dr. Dmitri Loguinov

Routing in sensor networks is a challenging issue due to inherent constraints such as power, memory, and CPU processing capabilities. In this thesis, we assume an *All to All* communication mode in an $N \times N$ grid sensor network. We explore routing algorithms which load balance the network without compromising the shortest paths constrain. We analyzed the *Servetto* method and studied two routing strategies, namely *Horizontal-Vertical* routing and *Zigzag* routing. The problem is divided into two scenarios, one being the static case (without failed nodes), and the other being the dynamic case (with failed nodes). In static network case, we derived mathematical formulae representing the *maximum* and *minimum* loads on a sensor grid, when specific routing strategies are employed. We show improvement in performance in load balancing of the grid by using *Horizontal-Vertical* method instead of the existing *Servetto* method. In the dynamic network scenario, we compare the performance of routing strategies with respect to probability of failure of nodes in the grid network. We derived the formulae for the *success-ratio*, in specific strategies, when nodes fail with a probability of p in a predefined source-destination pair communication. We show that the *Servetto* method does not perform well in both scenarios. In addition, *Hybrid* strategy proposed does not perform well compared to the studied strategies. We support the derived formulae and the performance of the routing strategies with extensive simulations.

To My Parents and Teachers

ACKNOWLEDGMENTS

"Give a man a fish a day, you feed him for a day. Teach him how to fish, you feed him for his life."-Anonymous

First, I sincerely thank my advisor Dr. Loguinov for providing me with my research problem. He has shown me the path to research and the critical aspects involved in research. He has instilled in me a state of confidence, with which I now feel that I have the ability to research any new topic following his research guidelines. He emphasized the importance of writing and expressing ones' research ideas. He has been a wonderful, inspiring presence in my life these past years, and I am forever in his debt. I also thank my committee members Dr. Klappenecker and Dr. Pappu for their support and insightful discussions about this work.

I would like to thank all the members of Internet Research Lab (IRL), TAMU for their help during this work. Especially, I would like to thank Derek for his help in LaTeX, and Xiaoming for his help in Excel charts. I thank my friends SaiKrishna Yadlapalli, Swaminathan Sethuraman and Kapil Bhattad for their critical discussions on my work. I would like to thank all those who have helped in reviewing this thesis.

Finally, I thank my parents and my family members for providing unending support in every aspect possible to ease my path on this journey. Without their support, I would not have been able to accomplish this task.

Last but not the least, I thank each and everyone, without whom this thesis would not have completed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

Sensor networks are deployed to monitor and provide feedback of environmental variables in areas, which are intractable to humans. With such deployments in mission critical applications, sensor networks gained importance and provide for immense potential for research in this area. Two challenging issues are identified in this realm. First, being the reduction in consumption of power by these sensors to increase their lifetime. Second, being the design of routing strategies for communication in the network.

In this thesis, we deal with routing strategies for *All to All(ATA)* communication in an $N \times N$ grid network in both, the static (without failed nodes), and the dynamic (with failed nodes) cases separately. We intend "dynamic" in a sense that the network might have node failure probability of p. However, we assume that once the *All-to-All* communication phase begins, no new nodes fail during the communication phase. Routing in sensor networks is a complex issue due to the large number of parameters. Unfortunately, there exists no single routing strategy which is considered to be efficient in all aspects. A routing strategy may be shown to be efficient based on obtaining minimum load on a particular node in the grid. However, that same strategy might not be efficient if if we consider load balance over all the nodes in the grid as the performance criterion. Therefore, it is better to decide the routing strategies based on the criteria of the application for which the sensor network is deployed. In addition, a routing strategy shown to be efficient in static network might not be an efficient strategy in dynamic networks. Hence, both the scenarios of

---

The journal model is *IEEE Transactions on Automation Control.*

static and dynamic networks are to be considered separately, and also the criterion for efficiency in both scenarios is to be decided.

In static network scenario, if we consider *success-ratio* as the performance criterion, then all the routing strategies would provide the same result and an efficient algorithm cannot be distinguished from an inefficient one. Here, we considered load balancing over all the nodes in the grid to be the criterion for deciding the efficiency of the routing algorithm. In dynamic network scenario, successful transmission of message from source to destination is more essential than load balancing over the nodes in the network. *Success-ratio*, defined as the fraction of messages that reach the destination successfully using shortest paths, is considered as the criterion for deciding the efficiency of the algorithm in dynamic case.

In static network case, we implemented *Horizontal-Vertical(H-V)*, *Zigzag*, and *Servetto* methods. *H-V* routing performs better than that of *Zigzag* and *Servetto* strategies, considering load balancing as the performance criterion. We modified the *H-V* method to obtain improved versions. We derived mathematical formulae representing the *maximum* and *minimum* load on the grid when specific routing methods are applied in the *All-to-All* communication mode.

In dynamic network case, we implemented the *H-V* approach, *Zigzag* method and *Hybrid* method (a combination of *H-V* and *Servetto* methods). We derived mathematical formulae for the probability of success of a path when routing strategies are applied in a single source–single destination mode.

This thesis is organized as follows: Section II discusses the sensor network applications; Section III deals with the background work; Section IV provides motivation and problem definition; Sections V, VI delve into routing strategies in static and dynamic network cases respectively; Section VII provides the conclusion of this thesis.

CHAPTER II

IMPORTANT APPLICATIONS OF SENSOR NETWORKS

Many of sensor network applications are discussed in [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. Due to their communication model, these networks have potential applications in many areas. Although the application areas are classified into environment, military and civil, these are not an exhaustive list of area of their application. Some of their applications are:

1. *Military Applications*: Sensors are widely used in applications such as surveillance, communication from intractable areas to base stations. Since these are inexpensive and deployed in large numbers, loss of some of these sensors would not affect the purpose for which they were deployed.

2. *Distributed Surveillance*: Highly mobile senor networks like the Underwater Autonomous vehicle Odessey make it possible to transmit huge amounts of data at low powers.

3. *WINS Wireless Sensing Networks*: These networks contain large arrays of distributed sensors and the interpolation (by making use of multiple sensors on each node) of various sensed datum gives high quality information. These networks are primarily used in military terrain and for monitoring complex machinery processes.

4. *Structure Monitoring*: Structure monitoring systems detect, localize, and estimate the extent of damage. Civil engineering structures can be tested for soundness using sensors.

5. *Pollution and Toxic Level Monitoring*: These sensors collect data from indus-

trial areas and areas where toxic spills occur. These are useful in sensing Nuclear, Biological, and Chemical phenomena in environment and transmitting it to remote stations for analysis.

6. *Sensors for Vision*: These are collaborative self-organizing sensor networks which have many micro sensors built on a chip and implanted in the eye. This improves the vision of people with ”no vision or limited vision”.

7. *Smart Sensor Networks*: These networks have a number of independent sensors. Each of the sensors makes a local decision and all the decisions are combined and weighed based on a specific algorithm and a global decision is taken.

8. *Rainfall and Flood Monitoring*: These networks have water level, wind and temperature sensors and the data is transmitted to a central database for analyzing and forecasting weather.

9. *Other Applications*: These involves habitat monitoring for determining bio-complexity. These include resource explorations such as mining and mineral analysis. Health applications involve tracking patients, monitoring drug administrations in hospitals. Great commercial opportunities exist in the household electronics and in realizing the smart home and office environments.

CHAPTER III

BACKGROUND AND RELATED WORK

Sensor networks, similar to mobile ad-hoc networks involve multi-hop communications. There have been many routing algorithms proposed for mobile networks. Yet, these algorithms are not applicable to sensor networks due to several factors as mentioned in [1]. Some of these factors are :

- The size of the sensor network is usually larger than that of ad-hoc networks. Sensor networks have high density of sensor nodes when compared to mobile hosts.
- Sensor nodes have energy constraints and are highly susceptible to failures. In addition, they are generally static.
- Sensor nodes use reverse multi-cast communication while ad-hoc networks use peer to peer communication.
- These nodes have several constraints with respect to power, memory, CPU processing which prohibits them from handling high data rate. Hence, sensors have low data rate than that of mobile hosts.

All these factors distinguish sensor networks from mobile networks, and make most of the routing protocols of mobile networks inapplicable to sensor networks. Hence, new routing algorithms are investigated for sensor networks. Routing in sensor networks is generally data centric [2]. The sensors sense specific data parameters and on querying about that parameter, they send their observations to the query initiator. Several papers [4], [6], [7], [8], [10] have addressed the issue of routing in sensor networks. However, many of these strategies are adaptive in behavior and not deterministic. Initial routing in sensor networks was done through flooding. The source node transmits message to all neighbors within its range. The corresponding neighbors recursively retransmit the message to their neighbors until the message reaches its destination.

Barett *et al.* in [4] developed an algorithm which would reduce the flooding in the network. In this paper [4], they reduce flooding by reducing the number of retransmissions. The nodes retransmit messages based on a probability function which depends on the distance of the node from the destination and number of times the message has been retransmitted. Instead of this kind of message flooding, Chalermek *et al.* proposed an algorithm based on data centric routing. This routing strategy in [7] is based on attribute-value querying and when queried, nodes establish gradients to the query initiator and send the attribute-value pair to the query initiating node. In [6], David *et al.* propose a refinement to the directed diffusion algorithm proposed in [7], named Rumor routing. Rumor routing is applicable in areas where nodes do not have a coordinate system. In this, the query generated is sent on randomly until it finds nodes which are on the path to the event destination. In another paper [11], Stefan *et al.* analyze the reliability of the system in the case of node failures. They split the data packet into multiple segments in such a way that the original data can be constructed from subset of all the segments. They route these mutliple segments on multiple paths and at the destination construct the original message from the messages received.

All the routing algorithms mentioned in [4], [6], [7], [8], [10], [12], [13] do not address the protocol performance in *All to All* communication mode. Although [14] deals with *All to All* communication model, it models the network topology as a tree rather than a mesh. Little research is done in finding routing algorithms which load balance the network. In [15], Goa *et al.* propose algorithms which utilize local information in routing and produce good load balance of the network. However, they assume that all the nodes in the network are arranged in a narrow strip whose width is constrained by the communication radii of the nodes by a small constant. In another paper [16], Dai *et al.* propose a load balancing algorithm on an asymmetric WSN topology having a tree rooted at base station to which the sensors communicate.

Servetto *et al.* recently proposed in [9], a routing algorithm (*Servetto*s' algorithm) which reduces the load on the central node in a single source–single destination communication. This algorithm divides the network into expansion and compression phases. Nodes belong to different diagonals of the grid. During expansion phase, the load per node decreases with the increase of number of nodes on diagonal. During the compression phase, the reverse process proceeds, and with the decrease in number of nodes on each diagonal, the load per node increases. Barrenechea *et al.* proposed an algorithm in [3], which performs better than that proposed in [9] and showed that in *All to All* communication mode, algorithm in [3] is optimal with respect to rate per node criterion. Barrenechea *et al.* proposed a hybrid approach which combines two existing routing strategies, and employs one of them based on the probability of failure of node in the grid.

Little research has been done in *Sensor* networks in investigating into algorithms which would perform well in *All to All* communication model. However, a lot of research has been done in the realm of *Computer Architecture*, in the context of *All to All* communication among processors. The processors are arranged in a mesh structure and every node sends and receives messages from every other node in the network. In [17], Sunggu et al. discuss the concept of *All to All* communication in meshes. However, their concern is to avoid link contention among transmitted messages and to reduce the time taken for this *All-to-All* communication phase. Another paper [18] by Rajeev *et al.* discuss certain algorithms for all to all communication. They consider link contention, time taken for each phase in an all to all communication mode and the minimum number of such phases required for the total exchange to take place. They deal with meshes whose size is not constrained to powers of two size meshes. In [19]. Susan *et al.* propose a hardware architecture for *All-to-All* communication. In another paper [20], Yang *et al.* analyze analytical models in *One-to-All* broadcast model and *All-to-All* strategies. The authors reduce the overall communication delay in *All-to-All* communication mode with the help of a pipelined approach, by overlapping the switching and transmission times of messages. In [21], Scott proved that the minimum number of contention-free steps for the *All-to-All* communication phase is $a^3/4$ in an $a \times a$ grid. However, they do not take load balancing criterion for routing. This shows that there is less work done in the area of deterministic load balancing routing algorithms in sensor networks.

CHAPTER IV

MOTIVATION AND PROBLEM DEFINITION

Routing in sensor networks has been a challenging issue for researchers considering the energy constraints in these networks. Deployment methodology also poses challenges in design of routing strategy. Sensors may be deployed deterministically or randomly based on the application for which they are used. For random applications, these sensors should be self configuring. These random deployments might result in irregular topologies which in turn affect the routing strategy. The preference of regular over an irregular topology simplifies the analysis phase without compromising the inherent constrains of the problem such as scalability [9]. Hence, we will consider a regular mesh topology for our problem. In these networks, messages are transmitted using multi-hop communications. Sensors perform both data sending and data routing. Inter-sensor communication is usually short ranged. The nodes in the network cooperate in forwarding other nodes' packets from source to destination. Hence, certain amount of energy of each node is spent in forwarding the messages of other nodes.

Usually, the central node will be heavily utilized in routing and forwarding messages, while the corner nodes are less utilized. This uneven load distribution results in heavily loaded nodes to discharge faster when compared to others. This causes few over-utilized nodes which fail and result in formation of holes in network, resulting in increase of *failed* messages in the network. A routing strategy developed should be such that it load balances the network and prevents the formation of holes. Servetto *et al.* in [9] proposed a spreading algorithm, known as the *Servetto* algorithm, which reduces the load over the central node and increases the load over the corner nodes in a single source–single destination communication. They claim that this routing

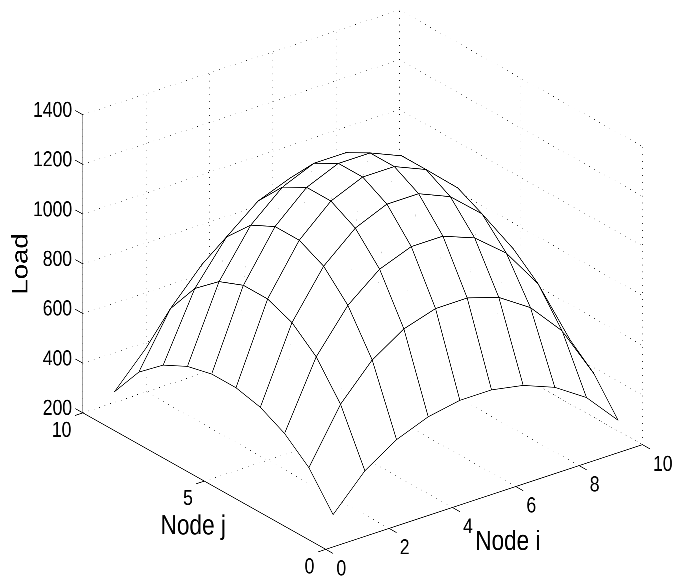


Fig. 1. Servetto Routing in 10×10 Grid (All to All Communication Mode)

strategy would result in minimizing the load over the center in such mesh networks. However, their results are applicable only when a single source–single destination communication is considered. When the communication is among all nodes, then the central node is observed to be the maximum load node as shown in Figure 1.

Little work is done in the area of developing routing algorithms which provide load balance of the network in an *All to All* communication scenario. One of the reasons for less investigation into load balancing routing algorithms is because, load balancing is considered to be an NP-hard problem in literature [15]. Load balancing routing and shortest path routing are conflicting features, as shortest path routing involves under utilization of some resources and load balancing routing requires utilization of all resources. Hence, there always exists a tradeoff between load balancing routing and shortest path routing. This motivated us to investigate routing algorithms which perform better load balance of the network using shortest paths. Algorithms are developed for both scenarios, for a static sensor network and for a dynamic network case. Extensive simulations are performed to support the inferences. We would provide the definitions, assumptions and notations used before proceeding further.

IV.A Definitions

Defintion 1. *A static Network is a network with no node failures.*

Defintion 2. *A dynamic Network is a network with failed nodes.*

Defintion 3. *All to All communication phase is a mode in which all the nodes in the network send and receive messages from all the other nodes in the network.*

Defintion 4. *Success-ratio is defined as the fraction of messages that reach their destination under node failure under shortest-path routing.*

The primary focus of this research is to delve into deterministic routing strategies which perform better load balancing of the mesh than existing strategies in an *All to All* communication scenario, using shortest paths. The following assumptions are made in this thesis:

IV.B Assumptions

Assumption 1. *We consider a regular mesh topology for sensor network as shown in Figure 2.*

This simplifies the analysis as well as modeling phases. Though simple, it considers some complex characteristics as scalability of the routing strategy with the size of the network.

Assumption 2. *We consider only shortest paths in routing from source to destination.*

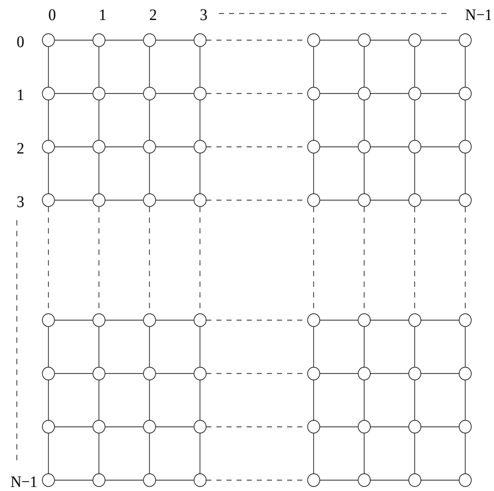


Fig. 2. Structure of $N \times N$ Grid

Assumption 3. *We consider the All to All communication mode.*

Assumption 4. *There are infinite-size buffers at each node of the grid to support the incoming and outgoing message packets.*

Hence,we do not consider buffer overflows and queueing analysis.

Assumption 5. *Once the All to All communication phase begins, no new nodes fail during the communication phase.*

Assumption 6. *Time taken for the All to All communication phase is not taken into consideration.*

IV.C Notations

Notations used in this thesis are mentioned in Table I.

Table I. Table of Notations

Symbol	*Representation Text*
(i, j)	Node on the i^{th} row and j^{th} column of the $N \times N$ grid
$L(i, j)$	Load on node (i, j) of the $N \times N$ grid
T	Total load on the $N \times N$ grid
p	Probability of failure of a node in the dynamic network grid
$M(i, j, k, l)$	Message sent from node (i, j) to node (k, l)
$RM(i, j)$	Messages routed through node (i, j)
$P(i, j, k, l)$	Probability of success of a path from node (i, j) to node (k, l)
L_{max}^{HV}	Maximum load on the grid when H-V routing is applied on the grid
L_{max}^{ZZ}	Maximum load on the grid when Zigzag routing is applied on the grid
L_{max}^{SV}	Maximum load on the grid when Servetto routing is applied on the grid
L_{min}^{HV}	Minimum load on the grid when H-V routing is applied on the grid
L_{min}^{ZZ}	Minimum load on the grid when Zigzag routing is applied on the grid
L_{min}^{SV}	Minimum load on the grid when Servetto routing is applied on the grid
L_{avg}	Average load on the grid
HVM-i	Horizontal-Vertical Method Variant - i

CHAPTER V

ROUTING STRATEGIES IN STATIC NETWORK SCENARIO

V.A Routing Strategies

This scenario considers a network without node failures. Two fundamental methods, *Horizontal-Vertical(H-V)* and *Zigzag*, have been explored and compared with *Servetto*'s method described in [9]. These strategies are applied in an *All to All* communication mode on an $N \times N$ grid network (shown in Figure 2) and their resultant load distributions are analyzed. The efficiency of the routing strategy is decided based on the load balancing criterion over all the nodes of $N \times N$ mesh. The performance of load balance is judged based on the maximum and minimum loads of the grid when different routing strategies are applied. Decreasing the maximum load and increasing the minimum load on the grid indicates improvement in performance of routing strategy. Figure 2 shows the grid structure considered. Nodes are denoted by small circles and packets are sent by every node to every other node in the grid. All interior nodes are assumed to have four-connectivity namely-an interior node (i, j) is connected to four other neighbor nodes namely $(i-1, j)$, $(i, j-1)$, $(i+1, j)$ and $(i, j+1)$. Now we look into the different routing strategies and their performance.

V.A.1 Horizontal-Vertical(H-V) Routing

H-V routing has been explored as XY routing in the field of computer architecture, in the context of inter-processor communication. The processors are assumed to be arranged in a mesh and messages are routed using this strategy among the processors [18]. However, in this context, researchers delved into the aspects of reducing the communication time among processors or considering the number of buffers required at each node for this communication. We use *H-V* routing and study its performance with respect to load balancing the network in an all to all communication mode. In this *H-V* method, if we consider (i, j) and (k, l) $(i \neq k,\ j \neq l)$as the source and the destination nodes respectively, then the message from source to destination is routed horizontally initially until the message reaches the node (i, l), thereafter it is routed in vertical direction to reach the destination node. Whenever there is a choice of path between horizontal and vertical directions, horizontal path is taken preference over vertical path.Throughout this section, *H-V* method is the routing strategy applied. The following features of *H-V* method are proved.

1. The load on any node (i, j) in the *horizontal-vertical* method is

$$L(i,j) \quad = \quad 2N\left[(N-i-1)\,i+(N-j-1)\,j\right]+(N-1)^2 \tag{5.1}$$

2. The total load on the nodes is

$$T \quad = \quad (2/3)\,N^3\left(N^2-1\right) \tag{5.2}$$

Lemma 1. *The load on any node (i, j) in the H-V method is equal to:*

$$L(i,j) \quad = \quad 2N\left[(N-i-1)\,i+(N-j-1)\,j\right]+(N-1)^2 \tag{5.3}$$

We observe that the above Equation 5.3 is symmetric in i and j. Therefore, node (i, j) has same load as the load on node $(N-i-1, N-j-1)$, $(N-i-1, j)$ and $(i, N-j-1)$. Equation 5.3 is obtained as shown below.

Proof. As shown in Figure 3, divide the whole region into 8 regions and consider the different load patterns flowing through node (i, j) as shown in Table II.

Table II. Table of Flow of Messages in *Horizontal-Vertical* Routing.

S. R.	*D. R.*	*Messages passing through* (i, j)
1	8	$ij\,(N-i-1)$
2	8	$i\,(N-i-1)\,(N-j-1)$
3	7	$(N-i-1)\,ji$
4	7	$(N-i-1)\,(N-j-1)\,i$
5	7	ji
5	2	$ji\,(N-j-1)$
5	6	$j\,(N-j-1)$
5	8	$j\,(N-i-1)$
5	4	$j\,(N-j-1)\,(N-i-1)$
6	1	$(N-i-1)\,ij$
6	7	$(N-j-1)\,i$
6	5	$(N-j-1)\,j$
6	3	$(N-j-1)\,(N-i-1)\,j$
6	8	$(N-j-1)\,(N-i-1)$
7	8	$i\,(N-i-1)$
8	7	$(N-i-1)\,i$

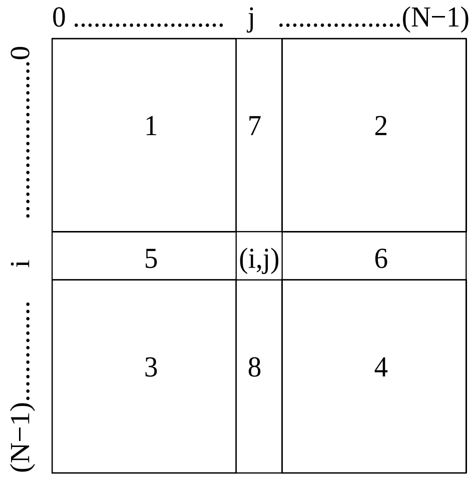


Fig. 3. Splitting Regions for Calculating Load on Node (i, j).

Now calculate the total load through node (i, j) due to other nodes in the system.Messages flowing through node (i, j) from all other regions to all other regions not mentioned in the Table II is zero.

$L(i, j)$= Summation of all messages mentioned in Table II.

$$= \quad 2N\left[(N-j-1)\,j+(N-i-1)\,i\right]+(N-1)^2 \tag{5.4}$$

□

Lemma 2. *The total load on the nodes is:*

$$T=(2/3)\,N^3\left(N^2-1\right) \tag{5.5}$$

Proof. From Equation 5.3, we have the load on (i, j):

$$L(i,j) \quad = \quad 2N\left[(N-j-1)\,j+(N-i-1)\,i\right]+(N-1)^2$$

Load on grid when nodes are only routing other nodes messages:

$$T' = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} L(i,j)$$

$$T' = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} 2N\left[(N-j-1)\,j + (N-i-1)\,i\right] + (N-1)^2$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} 2N\,(N-1)\,(j+i) + 2N\left(j^2+i^2\right) + (N-1)^2$$

$$= 2N^3\,(N-1)^2 + (N-1)^2\,N^2 + 2N^3\left(\frac{(N-1)\,(2N-1)}{3}\right)$$

$$= N^2\,(N-1)\left(\frac{6N^2-6N+3N-3-4N^2+2N}{3}\right)$$

$$T' = N^2\left(N^2-1\right)\left(\frac{2N-3}{3}\right)$$

The above scenario T' considers only the scenario in which all nodes perform forwarding of messages for other nodes in the network and ignores the messages for which these nodes are destination. Number of messages having a specific node as destination is (N^2-1), considering that node doesn't send to itself. Hence, for all nodes in the grid, it is $(N^2-1)\,N^2$. Total load on grid now can be obtained as:

$$\begin{aligned} T &= T' + (N^2-1)\,N^2 \\ &= N^2\,(N^2-1)\,[\,\left(\tfrac{2N-3}{3}\right)\,] + (N^2-1)\,N^2 \\ &= N^2\,(N^2-1)\,[\,\left(\tfrac{2N-3}{3}+1\right)\,] \\ T &= \tfrac{2}{3}N^3\,(N^2-1) \end{aligned} \tag{5.6}$$

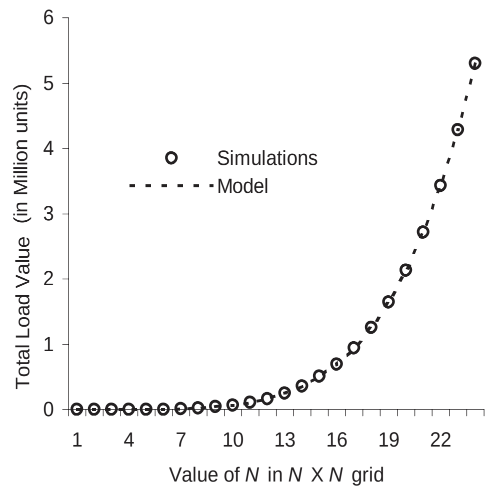


Fig. 4. Model for Total Load on Grid

□

This value above in Equation 5.6 is the same as the one derived previously for the total load in Equation 5.5 of the system. To confirm this Equation 5.6 through MATLAB by using the known loads, we consider the first few cases like 2 × 2, 3 × 3, 4 × 4 up to 10 × 10 nodes scenario and record the total load values for these cases. Figure 4 shows the total load distribution for different size grids. We find the equation of the curve that best fits these values. In this way, we derive at this Equation 5.6.

Corollary 1. *Average load on a node of the grid is:*

$$L_{avg} = \frac{2}{3}N\left(N^2 - 1\right) \tag{5.7}$$

Corollary 2. *Minimum load on the grid is:*

$$L_{min}^{VH} = (N-1)(3N+1) \tag{5.8}$$

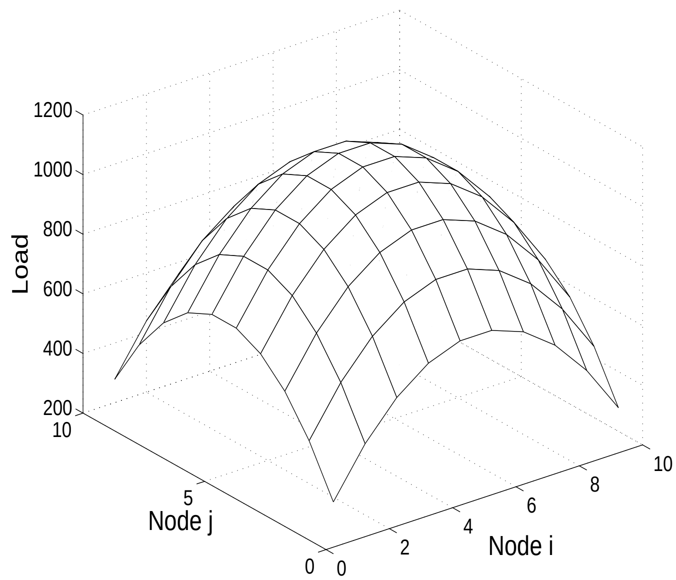


Fig. 5. Load on 10×10 Node Grid Using Horizontal-Vertical Routing.

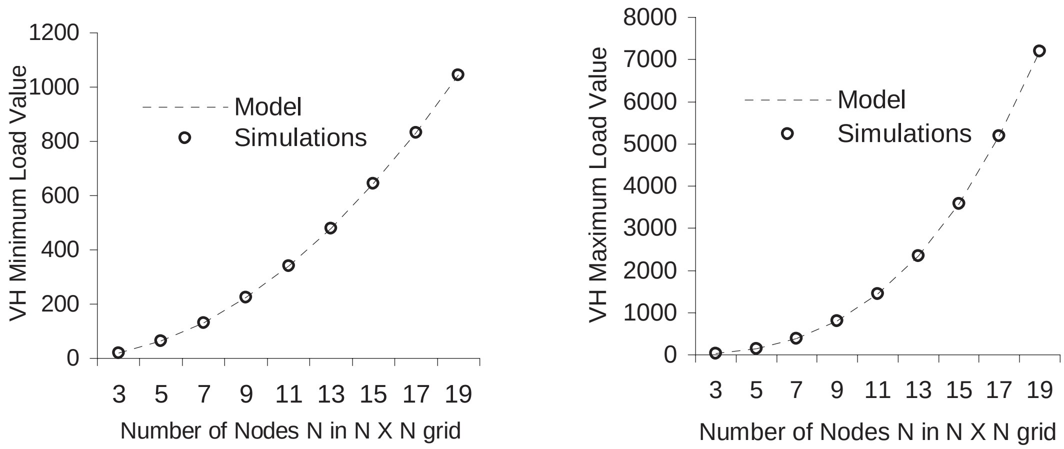
Corollary 3. *Maximum load on the grid is:*

$$L_{max}^{VH} = \begin{cases} (N+1)(N^2-1), & \text{if } N \text{ is odd} \\ (N+1)(N^2-2)+1, & \text{if } N \text{ is even} \end{cases}$$

Substituting the corresponding values of (i, j) of corner node and central node in Equation 5.3 yields minimum and maximum load of the grid respectively. Corner node has (i, j) as $(0, 0)$ correspondingly. Central node has different values depending on whether N is even or odd. When N is odd, then (i, j) values of central node are $(\frac{N-1}{2}, \frac{N-1}{2})$ correspondingly. When N is even, then (i, j) values are $(N/2, N/2)$ respectively.

Corollary 4. *Max/Min load ratio in Horizontal-Vertical routing is* $\Theta(\frac{N}{3})$.

Simulation: We have implemented *H-V* method and simulated it on a 10×10 grid. Figure 5 shows the load distribution produced from the simulation. This



(a) (b)

Fig. 6. Comparison of Load Models with Simulations (N is odd) (a)H-V Minimum Load Model, (b) H-V Maximum Load Model.

confirms that the central node has maximum load of the grid, and the corner nodes have the minimum load of the grid. It shows the visual symmetry in load distribution on the grid. Figure 4 confirms the total load formula derived in Equation 5.5. Figure 6 shows the result of comparing the model with the simulation values. Figures 6(a), 6(b) plot the simulation values against the derived model for odd values of N.

V.A.2 Zigzag Routing

In this method, messages are sent along a path which follows a zigzag pattern. When the message can no longer go in zigzag pattern, it reaches the destination using the horizontal or vertical path, which ever is appropriate. Some observations in *zigzag method* are:

1. In zigzag method, the corners are less utilized and the load on the corner nodes is equal to $(N-1)$ when we consider an $N \times N$ node grid.

2. Load on the boundary nodes is less and increases greatly as we move towards

the center.

3. Load is distributed symmetrically in horizontal and vertical directions. Hence, load on i^{th} node is same as the load on $(N-i-1)^{th}$ node in every column and load on j^{th} node is same as the load on $(N-j-1)^{th}$ node in every row.

4. The overall load in this method is equal to the overall load in the horizontal-vertical method.

5. At the boundaries, the load is less for zigzag method but as we move to the center the load in the zigzag method is larger than that of load in the horizontal-vertical method.

Lemma 3. *The total load on the nodes is:*

$$T = (2/3)\, N^3 \left(N^2 - 1\right) \tag{5.9}$$

Proof. Considering T as *Total Load* on the grid and $d(i,j,k,l)$ as the hop distance between node (i,j) and node (k,l) $(i \neq k,\ j \neq l)$, we have

$$d(i,j,k,l) = |i-k| + |j-l| \tag{5.10}$$

$$\begin{aligned}
T &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} d(i,j,k,l) \\
&= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \left(|i-k| + |j-l|\right) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \left(\ |i-k|\, N + N\left(\tfrac{N+1}{2}\right) - j\,(N+1) + j^2\ \right)
\end{aligned}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N} (N+1)(N-j-i)N + (j^2+i^2)N$$

$$= 2N^2\left[N(N+1)\left(\tfrac{2N+1}{6}\right)\right] + N^4(N+1) - N^3(N+1)^2$$

$$= \left(\tfrac{2}{3}\right)N^3(N^2-1)$$

Therefore total load on the grid

$$T = \left(\frac{2}{3}\right)N^3\left(N^2-1\right) \tag{5.11}$$

□

Additionally, we observe that the total load of *H-V* method (shown in Equation 5.5) is same as total load of *Zigzag* method (shown in Equation 5.9). Simulations of these strategies show that the central node is heavily loaded and the corner nodes are the least loaded nodes. Hence, it is sufficient if we compute the load on corner and central node to derive the minimum and maximum loads of the grid respectively.

Lemma 4. *Minimum load on a node of the grid is:*

$$L_{min}^{ZZ} = (N-1)(2N+3) \tag{5.12}$$

Proof. The corner nodes of the grid has minimum load on the grid. Considering the top left corner node (node $(0,0)$ as shown in Figure 2) Considering horizontal direction to be preferred over vertical direction, we know that only node $(0,1)$ will utilize node $(0,0)$ to route its packets destined for nodes aligned on the left edge. Hence in an $N \times N$ grid, (N-1) messages will be routed through $(0,0)$, going from $(0,1)$ to nodes on the same column as $(0,0)$. Considering the messages for which node $(0, 0)$ will be source and destination, the overall load on $(0, 0)$ is:

$$\begin{aligned} L^{ZZ}_{min} &= (N-1) + 2(N^2-1) \\ &= (N-1)(2N+3) \end{aligned}$$

□

Lemma 5. *Maximum load of a node of the grid is:*

$$L^{ZZ}_{max} = \begin{cases} \left(\frac{1}{2}\right)(3N+1)(N^2-1), & \text{if } N \text{ is odd} \\ \left(\frac{1}{4}\right)\left(6N^3 - N^2 - 2N - 12\right), & \text{if } N \text{ is even} \end{cases}$$

Proof. Divide the grid into eight regions (as shown in Figure 3). Since central node is the node which is heavily loaded compute the load on the center node. We have to consider the cases when N is odd and even. Table III shows the load routed through the center node due to messages sent from one region to the other region. Considering the case when N is odd, the number of messages sent from *Region* 1 or *Region* 2 or *Region* 3 or *Region* 4 to their destination regions result in equal load on the central node. Hence, calculating the number of messages sent from one region that are routed through central node is sufficient to calculate the total number of messages produced due to these four regions. Hence, summing up the other regions load shown in Table III to this we have:

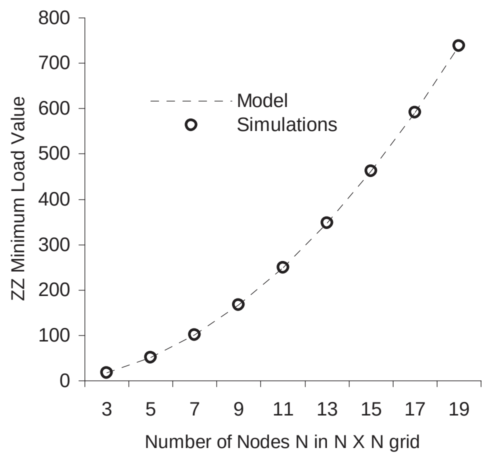
$$RM\left(\frac{n-1}{2}, \frac{n-1}{2}\right) = \left(\frac{3}{2}\right)(N-1)(N^2-1) \tag{5.13}$$

Messages for which the center node is source or destination is $2(N^2-1)$. Adding this value to Equation 5.13, we obtain the load on center node $(\frac{n-1}{2}, \frac{n-1}{2})$ as follows:

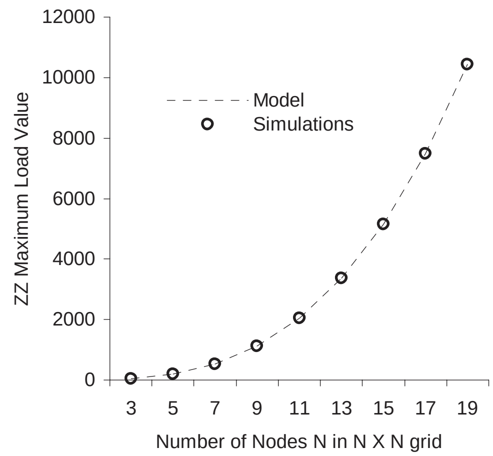
$$L\left(\frac{n-1}{2}, \frac{n-1}{2}\right) = \left(\frac{1}{2}\right)(3N+1)(N^2-1) \tag{5.14}$$

Table III. Table of Flow of Messages in *Zigzag* Routing.

S. R.	*D. R.*	*N is odd*	*N is even*
1	4	$(2i-1)(N-i-1)(N-j-1)$	$(2i-1)(N-i-1)(N-j-1)$
1	6	$i\left(\frac{i+1}{2}\right)(N-j-1)$	$i\left(\frac{i+1}{2}\right)(N-j-1)$
1	8	$\left(\frac{i^2+3i-2}{2}\right)(N-i-1)$	$\left(\frac{i^2+3i-2}{2}\right)(N-i-1)$
2	3	$(2i-1)(N-i-1)j$	$(2i-1)(N-i-1)j$
2	5	$i\left(\frac{i+1}{2}\right)j$	$i\left(\frac{i+3}{2}\right)j$
2	8	$\left(\frac{i^2+3i-2}{2}\right)(N-i-1)$	$\left(\frac{i^2+3i}{2}\right)(N-i-1)$
3	2	$(2i-1)(N-j-1)i$	$(2i-1)i(N-j-1)$
3	6	$i\left(\frac{i+1}{2}\right)(N-j-1)$	$i\left(\frac{i+1}{2}\right)(N-j-1)$
3	7	$\left(\frac{i^2+3i-2}{2}\right)i$	$\left(\frac{i^2+5i-2}{2}\right)i$
4	1	$(2i-1)(N-i-1)j$	$(2i+1)ij$
4	5	$i\left(\frac{i+1}{2}\right)j$	$(i+1)\left(\frac{i+2}{2}\right)j$
4	7	$\left(\frac{i^2+3i-2}{2}\right)i$	$\left(\frac{i^2+5i+2}{2}\right)i$
5	$2,4$	$(N-1)(N-j-1)$	$(N-1)(N-j-1)$
5	6	$j(N-j-1)$	$j(N-j-1)$
5	$7,8$	$N-1$	$N-1$
6	$1,3$	$(N-1)j$	$(N-1)j$
6	$7,8$	$N-1$	$N-1$
6	5	$j(N-j-1)$	$j(N-j-1)$
7	8	$(N-i-1)\,i$	$(N-i-1)\,i$
8	7	$(N-i-1)\,i$	$(N-i-1)\,i$



(a)



(b)

Fig. 7. Comparison of Load Models with Simulations (N is odd) (a)Zigzag Minimum Load Model, (b) Zigzag Maximum Load Model.

Since center node is the maximum loaded node we have as in Lemma 5. When N is even, we sum the messages in the Table III and obtain the value of load on the center node $(\frac{n-2}{2}, \frac{n-2}{2})$. $\square$

Corollary 5. *Minimum load of Zigzag routing is always lower than that of H-V routing.*

Since N is a positive integer we have $N > 0$ and $3N+1 > 2N+1$ always. Hence $(N-1)(3N+1) > (N-1)(2N+1)$. From this, Corollary 5 follows.

Corollary 6. *Max/Min load ratio in Zigzag routing is* $\Theta(\frac{3N}{4})$.

Simulation: As in *H-V* routing, the central node and corner node are the maximum and minimum loaded nodes. From the formulae derived, we observe that minimum load of *Zigzag* routing is always lower than the minimum load of *H-V* routing, and the maximum load of *Zigzag* routing is always higher than that of *H-V*

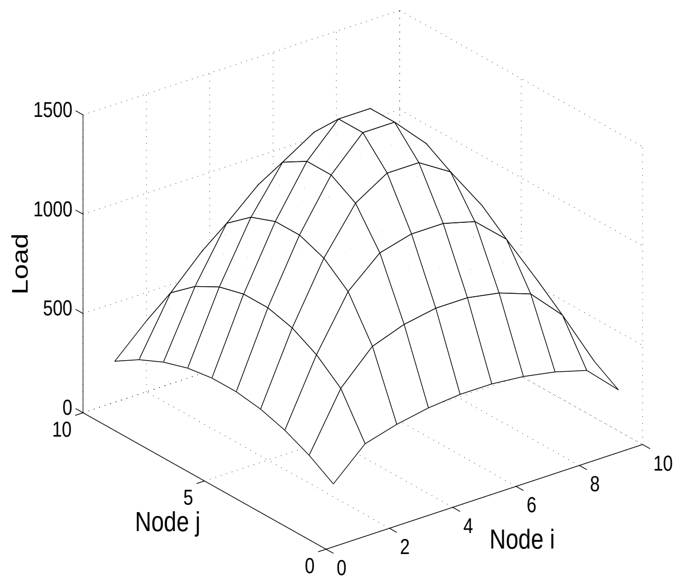


Fig. 8. Load on 10×10 Node Grid Using Zigzag Routing.

routing. This is confirmed from the simulations. The load value models are compared with the simulation values and results are shown in Figure 7. We implemented *Zigzag* routing and simulated the scenario on a 10×10 grid. Figure 8 shows the result of the simulation.

V.A.3 Servetto Method

Servetto method has been explored by Servetto *et al.* in [9]. In this, the nodes are considered to belong to specific diagonals on which they are present. In this method the load gets equally distributed on the nodes of the diagonal. Node (i, j) belongs to l^{th} diagonal if $(i + j)$ is equal to l. The whole network is divided into two stages comprising of expansion and compression phases. In expansion stage, load per node of the diagonals keeps decreasing and later in the compression stage the load per node of the diagonal increases as it goes towards the destination. In paper [9], Servetto *et al.* consider only a single source – single destination problem in which the source is at $(0,0)$ and the destination is at $(N-1,N-1)$ and show that the central node is a minimum loaded node. However, when we generate the all to all communication on the network, the central node has maximum load on the grid and has similar load distribution curve as others (*H-V*, *Zigzag*).

Lemma 6. *Minimum load on a node of the grid is:*

$$L_{min}^{SV} \approxeq 2\left[N^2+(2N+1)S_N-2N+1\right] \tag{5.15}$$

where S_N is given as:

$$S_N \approxeq \log\left(\frac{N}{\sqrt{2}}\right)-\frac{1}{12}+\frac{2}{3N}, N>2$$

Proof. The minimum load on the grid is at node $(0,0)$. Denoting $LM(i,j,k,l)$ as the load on node $(0,0)$ when a message is routed through $(0,0)$, sent from node (i,j) to node (k,l). Then messages routed through node $(0,0)$ denoted as $RM(0,0)$:

$$RM(0,0)=\sum_{x=1}^{N-1}\sum_{y=1}^{N-1} LM(x,0,0,y)+LM(0,y,x,0) \tag{5.16}$$

In *Servetto* algorithm we have:

$$\forall x,y : LM(x,0,0,y)=LM(0,y,x,0) \tag{5.17}$$

From Equations 5.16, 5.17 we have:

$$\begin{aligned} RM(0,0) &= 2\sum_{x=1}^{N-1}\sum_{y=1}^{N-1} LM(x,0,0,y) \\ &= 2\sum_{k=2}^{N}\frac{2N-2k+1}{k} \\ &= 2\left[(2N+1)\left(\sum_{k=2}^{N}\frac{1}{k}\right)-2(N-1)\right] \\ &= 2\left[(2N+1)S_N-2(N-1)\right] \end{aligned} \tag{5.18}$$

The integral value can be calculated using Simpsons' rule as below. We can assume N to be even without loss of generality.

$$\begin{aligned}
\int_{k=1}^{N} \frac{1}{k} dk &= \frac{1}{3}\left[1 + \frac{1}{N} + 4\left(\frac{1}{2} + \frac{1}{4}\cdots\frac{1}{N-2}\right) + 2\left(\frac{1}{3} + \frac{1}{5}\cdots + \frac{1}{N-1}\right)\right] \\
\int_{k=2}^{N} \frac{1}{k} dk &= \frac{1}{3}\left[\frac{1}{2} + \frac{1}{N} + 2\left(\frac{1}{4}\cdots\frac{1}{N-2}\right) + 4\left(\frac{1}{3} + \frac{1}{5}\cdots + \frac{1}{N-1}\right)\right]
\end{aligned} \tag{5.19}$$

Adding the two equations in Equation 5.19, we have:

$$\begin{aligned}
\int_{k=1}^{N} \frac{1}{k} dk + \int_{k=2}^{N} \frac{1}{k} dk &= \frac{1}{3}\left[\frac{1}{2} + 6S_N - \frac{4}{N}\right], N > 2 \\
\log(N) + \log\left(\frac{N}{2}\right) &= \frac{1}{6} + 2S_N - \frac{4}{3N}, N > 2 \\
\log\left(\frac{N^2}{2}\right) &= \frac{1}{6} + 2S_N - \frac{4}{3N}, N > 2
\end{aligned} \tag{5.20}$$

Now rearranging terms in Equation 5.20, we have:

$$S_N \approxeq \log\left(\frac{N}{\sqrt{2}}\right) - \frac{1}{12} + \frac{2}{3N}, N > 2 \tag{5.21}$$

Number of messages for which $(0, 0)$ is the source or destination is $2(N^2 - 1)$. Hence, total load on node $(0, 0)$ is:

$$\begin{aligned}
L(0,0) &= RM(0,0) + 2(N^2 - 1) \\
&= 2\left[(2N+1)S_N - 2(N-1)\right] + 2(N^2 - 1) \\
&= 2\left[N^2 + (2N+1)S_N - 2N + 1\right]
\end{aligned} \tag{5.22}$$

where S_N is given as:

$$S_N \approxeq \log\left(\frac{N}{\sqrt{2}}\right) - \frac{1}{12} + \frac{2}{3N}, N > 2$$

Since node $(0,0)$ is the minimum load bearing node of the grid, This should be the minimum load of the grid. As we have approximated the value of $RM(0,0)$, there would be a slight deviation in the obtained value. $\square$

Lemma 7. *Maximum load on a node of the grid when N is odd is:*

$$\begin{aligned} &= 4\left[(2k+3)^3\log(2k+3)\right] - 8\left[(4k^3+21k^2+27k)\log(k+3)\right] \\ &+ 8(2k+1)(2k+3)S_{k+1} + 12(2k^2-9)\log(3) - \left(\frac{44k^3}{3}\right) - 68k^2 - 80k \end{aligned} \tag{5.23}$$

where S_{k+1} is given as:

$$S_{k+1} \approxeq \log\left(\frac{k+1}{\sqrt{2}}\right) - \frac{1}{12} + \frac{2}{3(k+1)}, k > 1$$

Proof. We consider an odd value of N so that the derivation phase is comprehensible and easier. As shown in Figure 9, divide the whole region into eight parts and then compute the load on the central node due to messages sent among these regions. Assuming $RM(R_i, R_j)$ denoting the load on central node when messages are sent from region i to region j, we have

$$RM(R_i, R_j) = RM(R_j, R_i) \tag{5.24}$$

and we also see that due to symmetry, we have the following equations:

$$RM(R_1, R_4) = RM(R_2, R_3) \text{ and } RM(R_5, R_6) = RM(R_7, R_8) \tag{5.25}$$

$$RM(R_5, R_7) = RM(R_5, R_8) = RM(R_6, R_7) = RM(R_6, R_8) \tag{5.26}$$

$$\begin{aligned} RM(R_1, R_6) &= RM(R_1, R_8) = RM(R_2, R_5) = RM(R_2, R_8) \\ &= RM(R_3, R_7) = RM(R_3, R_6) = RM(R_4, R_5) = RM(R_4, R_7) \end{aligned} \tag{5.27}$$

Hence, from all these equation above, we observe that we are required to compute

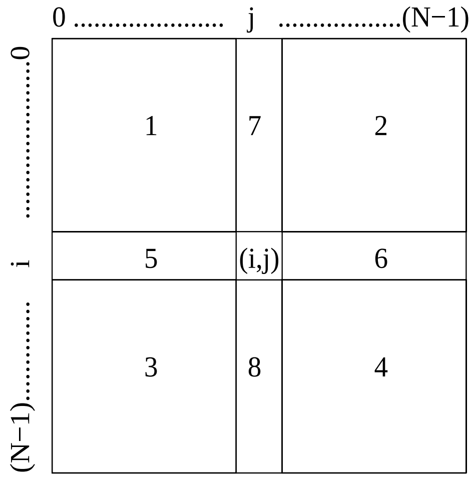


Fig. 9. Splitting Regions for Calculating Load on Central Node (i, j) in Servetto Routing.

only four of these quantities, namely - $RM(R_1, R_4)$, $RM(R_1, R_6)$, $RM(R_5, R_6)$, and $RM(R_5, R_7)$ respectively. Assuming the center of the $N \times N$ grid to be (k, k):

$$RM(R_5, R_6) = k^2 \tag{5.28}$$

$$\begin{aligned} RM(R_5, R_7) &= \left[\left(\frac{2k-1}{2}\right) + \left(\frac{2k-3}{3}\right) \cdots \left(\frac{1}{k+1}\right)\right] && (5.29) \\ &= (2k+3)\left(\sum_{j=2}^{k+1}\left(\frac{1}{j}\right)\right) - 2k \\ &= (2k+3)\, S_{k+1} - 2k && (5.30) \end{aligned}$$

We have dealt with a similar equation before in Equation 5.18 when calculating the minimum load of $N \times N$ grid using *Servetto* routing. Hence, we can utilize that calculated value here. This is because the central node of $N \times N$ grid will become the right bottom corner node in the $(k+1) \times (k+1)$ grid. Hence calculating the $RM(R_5, R_7) + RM(R_7, R_5)$ is equivalent to calculating the $RM(0, 0)$ value on a $(k+1) \times (k+1)$ grid. Now consider $RM(R_1, R_6)$ regions and the load due to this on the center is given as follows:

$$\begin{aligned} RM(R_1, R_6) &= k\left[\left(\frac{2k-1}{2}\right) + \left(\frac{2k-3}{3}\right)\cdots\left(\frac{1}{k+1}\right)\right] \\ &= k\left[(2k+3)\,S_{k+1} - 2k\right] \text{ from Equation 5.30} \\ &= k\left[(2k+3)\,S_{k+1} - 2k\right] \quad (5.31) \end{aligned}$$

Now considering $RM(R_1, R_4)$, we have:

$$\begin{aligned} RM(R_1, R_4) &= \sum_{a=0}^{k-1}\sum_{b=0}^{k-1}\sum_{c=k+1}^{N-1}\sum_{d=k+1}^{N-1}\left(\frac{1}{\min(a+b-2k, 2k-c-d)+1}\right) \quad (5.32) \\ &= \sum_{x=0}^{k-1}\sum_{y=0}^{k-1}\sum_{z=0}^{k-1}\sum_{w=0}^{k-1}\left(\frac{1}{\min(x+y, z+w)+3}\right) \quad (5.33) \\ &= \int_0^k\int_0^k\int_0^k\int_0^k\left(\frac{1}{\min(x+y, z+w)+3}\right) dwdzdydx \\ & \quad (5.34) \end{aligned}$$

However, we have observed that this Equation 5.34 is equivalent to a double integral with a modified function associated with certain probability density functions. We define $P = x+y$ and $Q = z+w$. Then we can write Equation 5.34 as $k^4 E(\frac{1}{min(P,Q)+3})$.

$$\begin{aligned} \int_{x=0}^{k}\int_{y=0}^{k}\int_{z=0}^{k}\int_{w=0}^{k}\left(\frac{1}{\min(x+y, z+w)+3}\right) dwdzdydx &= \quad (5.35) \\ \underbrace{k^4\int_{p=0}^{2k}\int_{q=0}^{2k}\left(\frac{1}{\min(p,q)+3}\right) g(p)g(q)dqdp}_{I} & \quad (5.36) \\ & \quad (5.37) \end{aligned}$$

where $g(x)$ is a probability density function defined as below:

$$g(x) = \begin{cases} \left(\frac{1}{k^2}\right)x, 0 \le x \le k \\ \left(\frac{1}{k^2}\right)(2k-x), k < x \le 2k \end{cases}$$

Equation 5.36 is verified by the simulations and the result is plotted in Figure 10.

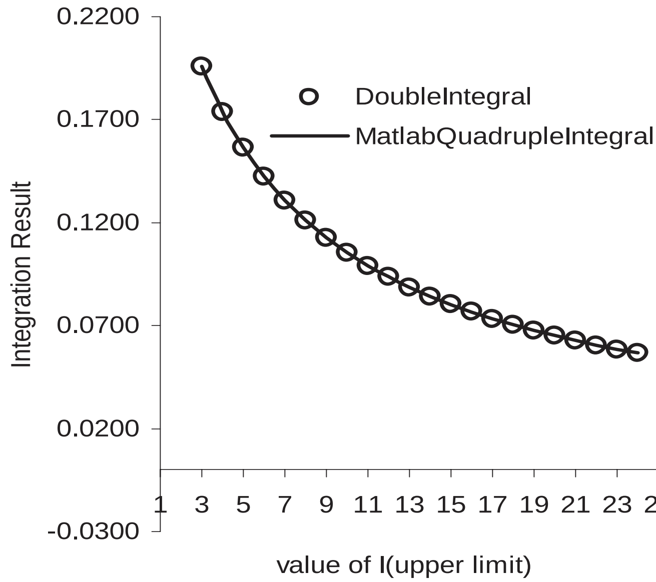


Fig. 10. Comparison of Double and Quadruple Integrals.

Now splitting Equation 5.36 into its constituent simpler integrals we have:

$$I = \underbrace{k^4 \int_{p=0}^{2k} \int_{q=0}^{p} \frac{1}{q+3} g(p)g(q)dqdp}_{I_1} + \underbrace{k^4 \int_{p=0}^{2k} \int_{q=p+1}^{2k} \frac{1}{p+3} g(p)g(q)dqdp}_{I_2} \tag{5.38}$$

$$\begin{aligned} I_1 = & \underbrace{\int_{p=0}^{k} \int_{q=0}^{p} \frac{1}{q+3} pqdqdp}_{I_{1,1}} + \underbrace{\int_{p=k}^{2k} \int_{q=0}^{k} \frac{1}{q+3} q(2k-p)dqdp}_{I_{1,2}} \\ & + \underbrace{\int_{p=k}^{2k} \int_{q=k}^{p} \frac{1}{q+3} (2k-q)(2k-p)dqdp}_{I_{1,3}} \end{aligned} \tag{5.39}$$

$$\begin{aligned} I_1 = & \underbrace{\int_{p=0}^{k} \int_{q=p}^{k} \frac{1}{p+3} pqdqdp}_{I_{2,1}} + \underbrace{\int_{p=0}^{k} \int_{q=k}^{2k} \frac{1}{p+3} p(2k-q)dqdp}_{I_{2,2}} \\ & + \underbrace{\int_{p=k}^{2k} \int_{q=p}^{2k} \frac{1}{p+3} (2k-q)(2k-p)dqdp}_{I_{2,3}} \end{aligned} \tag{5.40}$$

Solutions of these integrals are mentioned below:

$$I_{1,1} = \frac{k^3}{3} - \frac{3}{2}(k^2-9)\left[\log(k+3)-\log(3)\right] + \frac{3k^2}{4} - \frac{9k}{2} \quad (5.41)$$

$$I_{1,2} = \frac{k^3}{2} - \frac{3k^2}{2}\left[\log(k+3)-\log(3)\right] \quad (5.42)$$

$$I_{1,3} = \frac{-16k^3}{3} - \frac{27k^2}{4} - \frac{9k}{2} + \frac{(2k+3)^3}{2}\left[\log(2k+3)-\log(k+3)\right] \quad (5.43)$$

$$I_{2,1} = \frac{k^3}{3} - \frac{3}{2}(k^2-9)\left[\log(k+3)-\log(3)\right] + \frac{3k^2}{4} - \frac{9k}{2} \quad (5.44)$$

$$I_{2,2} = \frac{k^3}{2} - \frac{3k^2}{2}\left[\log(k+3)-\log(3)\right] \quad (5.45)$$

$$I_{2,3} = \frac{-16k^3}{3} - \frac{27k^2}{4} - \frac{9k}{2} + \frac{(2k+3)^3}{2}\left[\log(2k+3)-\log(k+3)\right] \quad (5.46)$$

As we see from Equations 5.41 – 5.46, the values of $I_{1,1}$, $I_{1,2}$, $I_{1,3}$ are equal to values of $I_{2,1}$, $I_{2,2}$, $I_{2,3}$ respectively. Finally we have:

$$\begin{aligned} I &= I_{1,1} + I_{1,2} + I_{1,3} + I_{2,1} + I_{2,2} + I_{2,3} \\ &= (2k+3)^3\log(2k+3) - 2(4k^3+21k^2+27k)\log(k+3) \\ &+ 3(2k^2-9)\log(3) - \frac{11k^3}{3} - 12k^2 - 18k \end{aligned} \quad (5.47)$$

From Equations 5.34, 5.36 we have

$$\begin{aligned} RM(R_1,R_4) = (2k+3)^3\log(2k+3) - 2(4k^3+21k^2+27k)\log(k+3) \\ +3(2k^2-9)\log(3) - \frac{11k^3}{3} - 12k^2 - 18k \end{aligned} \quad (5.48)$$

Load on center $RM(k,k)$ due to routing of messages sent by other nodes in the grid:

$$\begin{aligned}
RM(k,k) &= 4RM(R_1,R_4) + 4RM(R_5,R_6) + 8RM(R_5,R_7) + 16RM(R_1,R_6) \\
&= 4\left[(2k+3)^3\log(2k+3) - 2(4k^3+21k^2+27k)\log(k+3)\right] \\
&+ 4\left[3(2k^2-9)\log(3) - \frac{11k^3}{3} - 12k^2 - 18k\right] + 4\left[k^2\right] \\
&+ 8(2k+1)\left[(2k+3)\,S_{k+1} - 2k\right] \\
&= 4\left[(2k+3)^3\log(2k+3)\right] - 8\left[(4k^3+21k^2+27k)\log(k+3)\right] \\
&+ 8(2k+1)(2k+3)S_{k+1} + 12(2k^2-9)\log(3) - \left(\frac{44k^3}{3}\right) - 76k^2 - 88k
\end{aligned} \tag{5.49}$$

Since Maximum load on the grid is equal to load on center node, maximum load is:

$$\begin{aligned}
L(k,k) &= RM(k,k) + 2(N^2-1) \\
&= RM(k,k) + 2((2k+1)^2-1) \\
&= RM(k,k) + 2(4k^2+4k) \\
&= 4\left[(2k+3)^3\log(2k+3)\right] - 8\left[(4k^3+21k^2+27k)\log(k+3)\right] \\
&+ 8(2k+1)(2k+3)S_{k+1} + 12(2k^2-9)\log(3) - \left(\frac{44k^3}{3}\right) - 68k^2 - 80k
\end{aligned} \tag{5.50}$$

where S_{k+1} is given as:

$$S_{k+1} \approxeq \log\left(\frac{k+1}{\sqrt{2}}\right) - \frac{1}{12} + \frac{2}{3(k+1)}, k > 1$$

□

Corollary 7. *Max/Min load ratio in Servetto routing is* $\Theta(0.939N)$.

Simulation: We have implemented *Servetto* method and simulated it on a 10×10 grid. Figures 11, 12 show the load distribution on the grid when *Servetto*

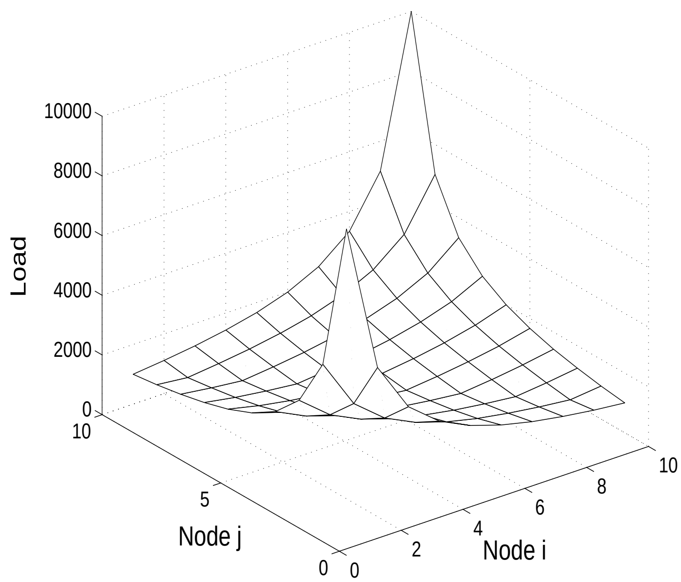


Fig. 11. Load on 10×10 Node Grid Using Servetto Routing on Single Source-destination Pair.

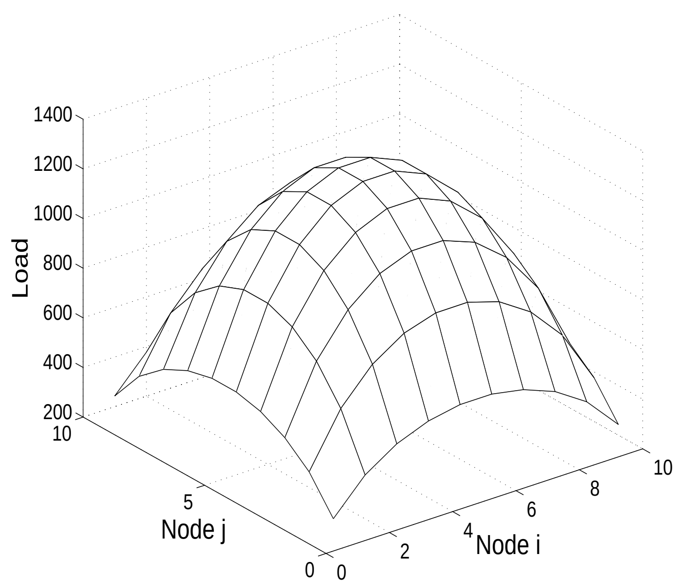
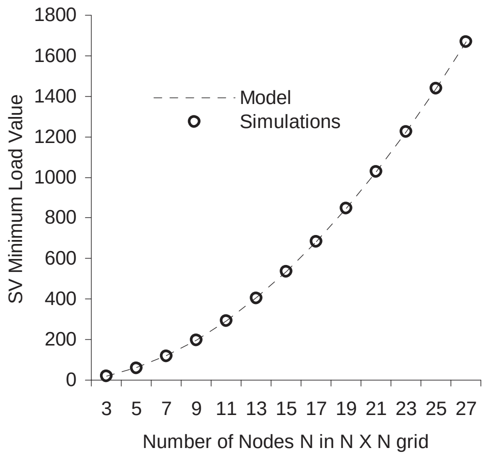
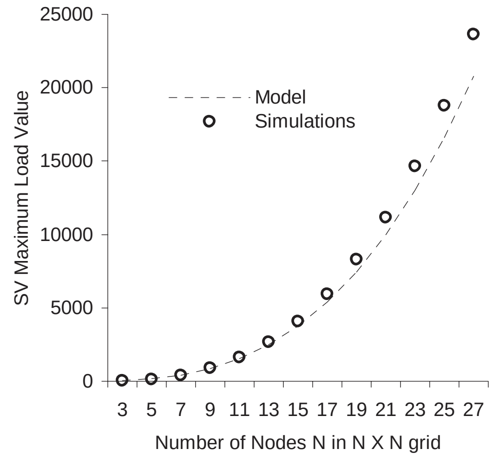


Fig. 12. Load on 10×10 Node Grid Using Servetto Routing.



(a)



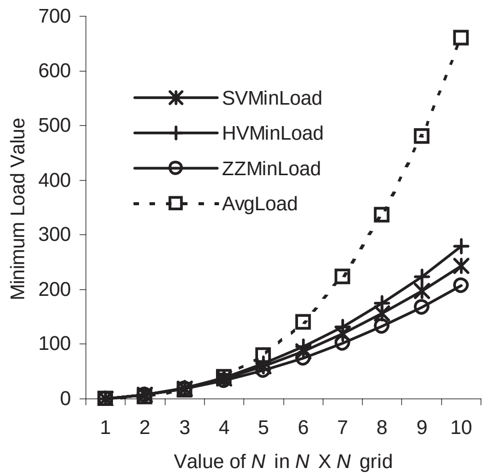
(b)

Fig. 13. Comparison of Load Models with Simulations (N is odd) (a)Servetto Minimum Load Model, (b) Servetto Maximum Load Model.

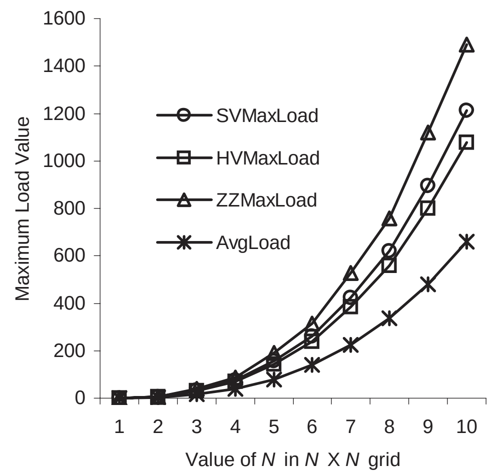
strategy is applied in a single source-single destination scenario and in all to all communication scenario respectively. The simulations confirm that the central node has maximum load of the grid and the corner nodes have the minimum load of the grid. In addition, simulation results show visual symmetry of load on nodes verifying the derived formulae.Though *Servetto* method applied for a single source – single destination reduces the load over the center (shown in Figure 11), *Servetto* method's performance in the *All to all* communication scenario (shown in Figure 12) is lower than that of *H-V* method (shown in Figure 5). We have compared the simulated values with the models generated as shown in Figure 13. The discrepancy in the two values is a result of approximation of summations to integration.

V.A.4 Comparison of Routing Strategies

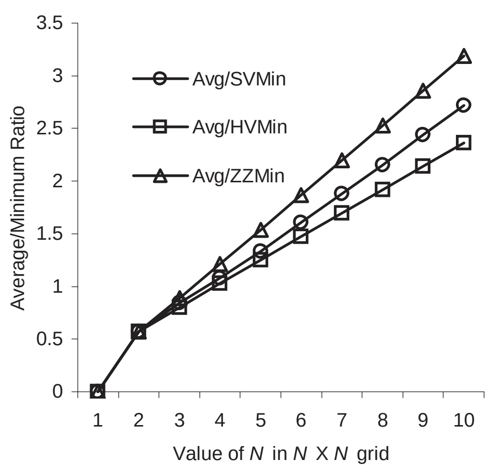
Here, we would compare the three strategies namely - *H-V*, *Zigzag* and *Servetto* routing implemented and the simulations are conducted for different values of N. The graphs have been plotted for various simulations with value of 'N' set to 10 and in these plots, the horizontal plane represents the coordinates of the node on the grid and 'z' axis represents the load on that node. The results of the simulations are shown below in Figure 14(a) – Figure 14(d). Figure 14(a) shows that the minimum load value in *Horizontal-Vertical* method is larger than the minimum load value of *Servetto*, and *Zigzag* methods for all values of N respectively. Figure 14(b) shows that the maximum load value for the *H-V* method is smaller than the maximum load value in other methods. Figure 14(d), Figure 14(c) show respectively that the *H-V* method has lower maximum/average load ratio as well as lower average/minimum load ratio for all values of N. Therefore *H-V* method performs better load balancing of the grid than the other two methods discussed namely *Zigzag* and *Servetto*.



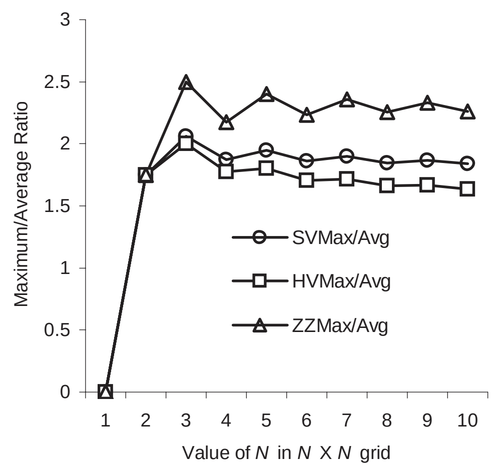
(a)



(b)



(c)



(d)

Fig. 14. Comparison of All Methods (a) for Minimum Load, (b) for Maximum Load, (c) for Average/Minimum Load Ratio, (d) for Maximum/Average Load Ratio.

V.B Refined Routing Strategies

V.B.1 Variants of Horizontal-Vertical Routing

With the result that the *Horizontal-Vertical* method has performed better than the *Servetto* and *Zigzag* methods, we delve more in improving these strategies by simulating several variations of the *Horizontal-Vertical* method with certain restrictions. In all these below described methods, consider a node (i, j) sending message to node (k, l) and the simulations are performed on a 10×10 grid. The different variants of horizontal-vertical method simulated:

1. *Without Center:* In this method, message routing is done in such a way that the center is avoided to be on the path from the source to destination. In cases where center does not lie on the paths to destination, then the path which is farther from the center is chosen. Simulation results are shown in Figure 15. Here we observe that though the load on the center is reduced, that decrease of

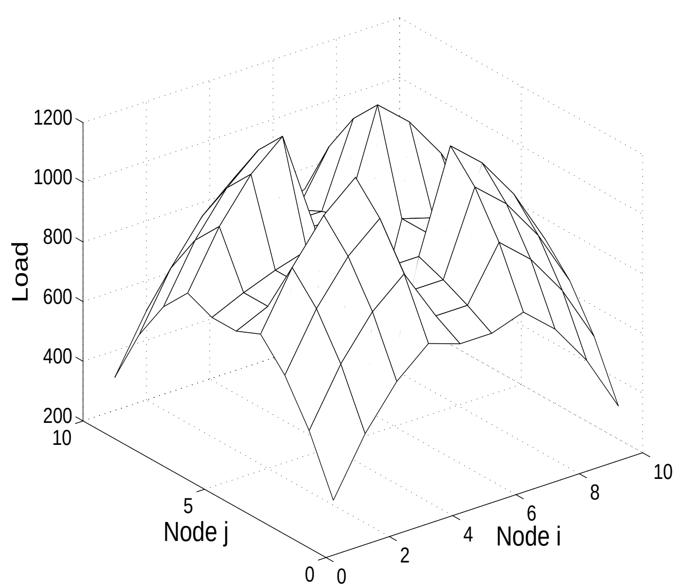


Fig. 15. Load on 10×10 Grid Using Horizontal-Vertical Method Variant-1

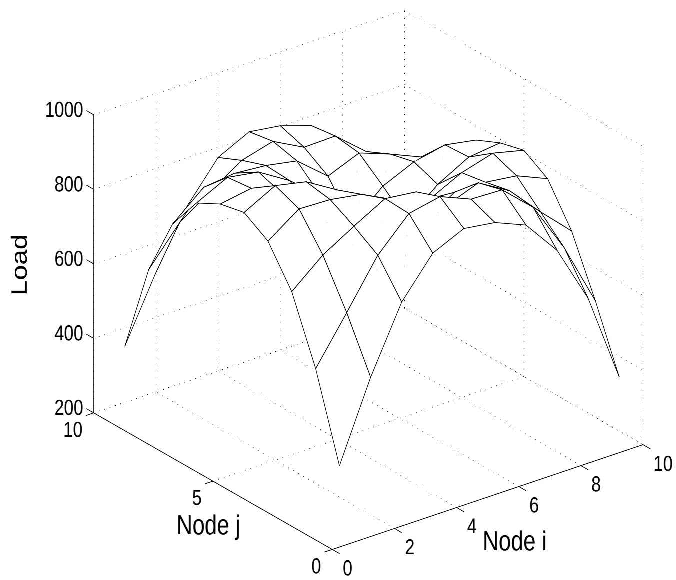


Fig. 16. Load on 10 × 10 Grid Using Horizontal-Vertical Method Variant-2

load on the central node is compensated with an increase of load on the nodes around the center.

2. *Avoiding Center:* In this strategy, the row distance $(i - x)$ and the column distance $(j - y)$ of the node (i, j) from the center (x, y) is calculated and the path having more distance from the center is taken. If the row distance is larger than or equal to column distance, then the horizontal path is taken, otherwise the vertical path is taken. If the center is one of the nodes on the path to the destination, then the above rule is violated and the alternative path is taken. Figure 16 shows the load distribution.

3. *Actual Distance From Center:* The actual distance from the center is calculated using the co-ordinates of the nodes. At every step, there are two choices for the selection of the next node on the path to destination. The actual distance of these two nodes from the center is calculated and the node which is more distant from the center is chosen as the next node on the path to its destination.

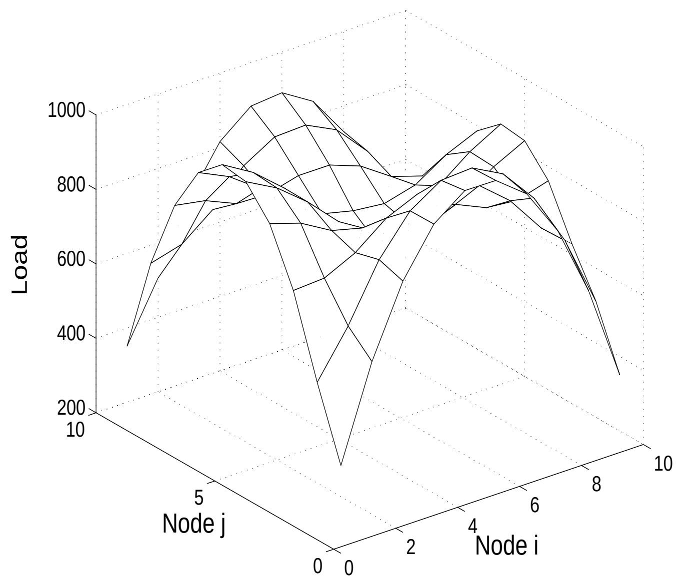


Fig. 17. Load on 10×10 Grid Using Horizontal-Vertical Method Variant-3

Repeatedly, this same procedure is applied at every node until it reaches its destination. Figures 15, 16, 17 provide the plots of load distribution of the horizontal vertical method variations - 1,2,3 respectively.

4. *Average Distance From Center:* In this strategy, all the nodes on the path to the destination are taken and their average distance from the center , along the two available paths from source to destination is calculated separately. Then the path having larger average distance from the center is selected for routing the message from source to destination. Figure 18 shows the load distribution on the grid.

5. *Minimum Distance From Center:* In this method, all the nodes on the path from source to destination are considered and their distances from the center are measured. Along the two possible paths from the source to destination, the minimum distance of the considered nodes is calculated separately. Then the message is routed through the path which has larger minimum distance of

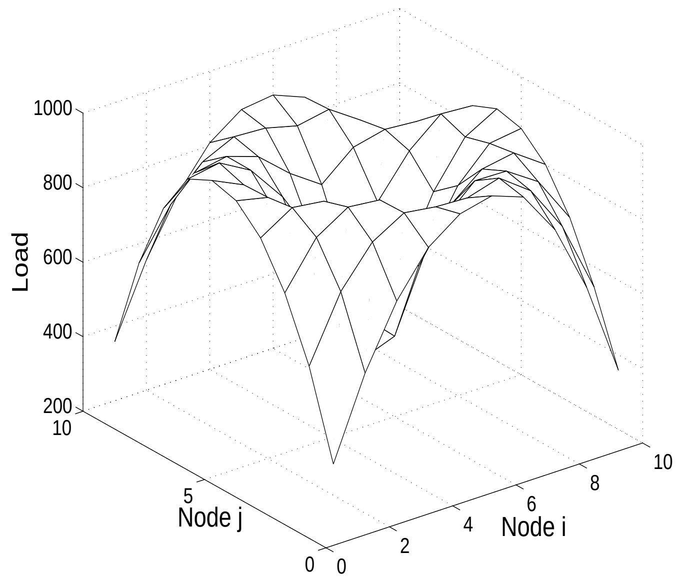


Fig. 18. Load on 10 × 10 Grid Using Horizontal-Vertical Method Variant-4

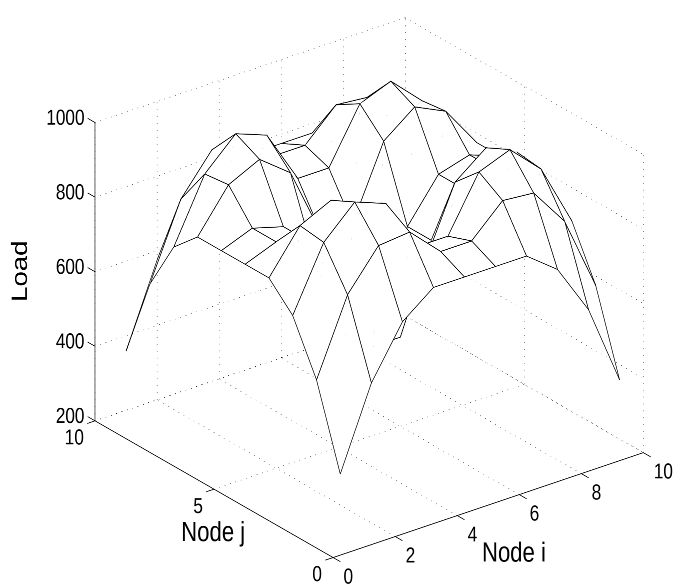


Fig. 19. Load on 10 × 10 Grid Using Horizontal-Vertical Method Variant-5

the two paths considered. This scenario inherently avoids the situation of the center being as a node on the path from source to destination. Figure 19 shows the load distribution.

V.B.2 Comparison of Variant Strategies

First the horizontal-vertical method variations–1,2,3 are simulated. The graphs are as shown in Figure 20 and we observe that the method variation–2 performs better than the other variations of horizontal–vertical methods. Then the methods–4,5 are simulated and compared with method 2, which was found to be better performing than all other previous methods-1,3. Figure 21(b) shows that variation method-2 has lower maximum load value when compared to other method variations-4,5. However Figure 21(a) does not indicate specifically which one performs better due to different methods performing better with different values of N. Therefore, there was no clear indication as to which method of these (variations -2,4,5) performed better. Yet, method variation-5 is expected to perform better and maybe with slight change in the implementation of the algorithm would show clear indication of its better performance when compared with other variations. The fluctuations show that maybe the algorithm needs to be manipulated differently based on values of N, specifically taking into consideration whether N is odd or even.

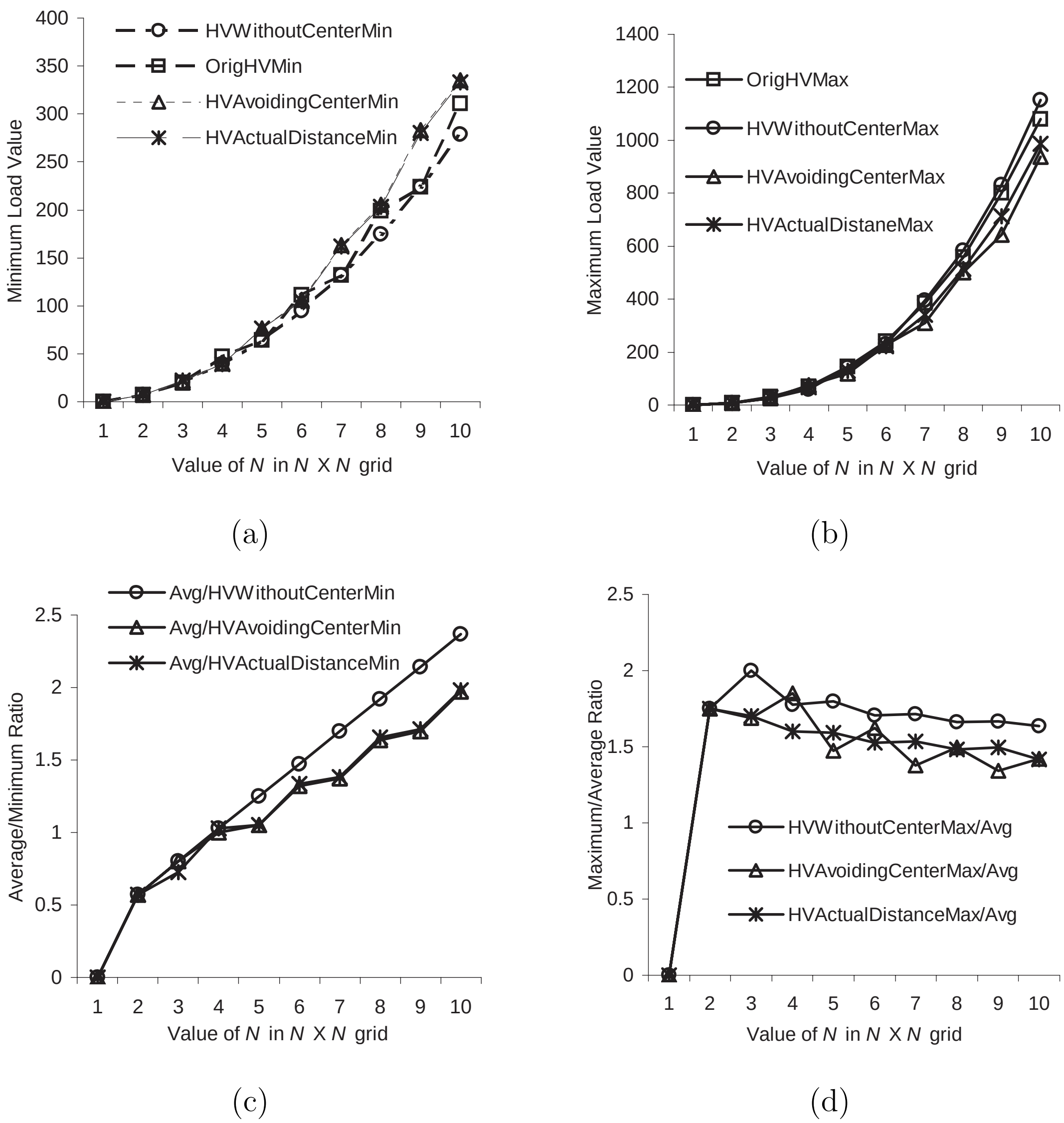


Fig. 20. Comparison of HVM-1,2,3 Variants (a) for Minimum Load, (b) for Maximum Load, (c) for Average/Minimum Load Ratio, (d) for Maximum/Average Load Ratio.

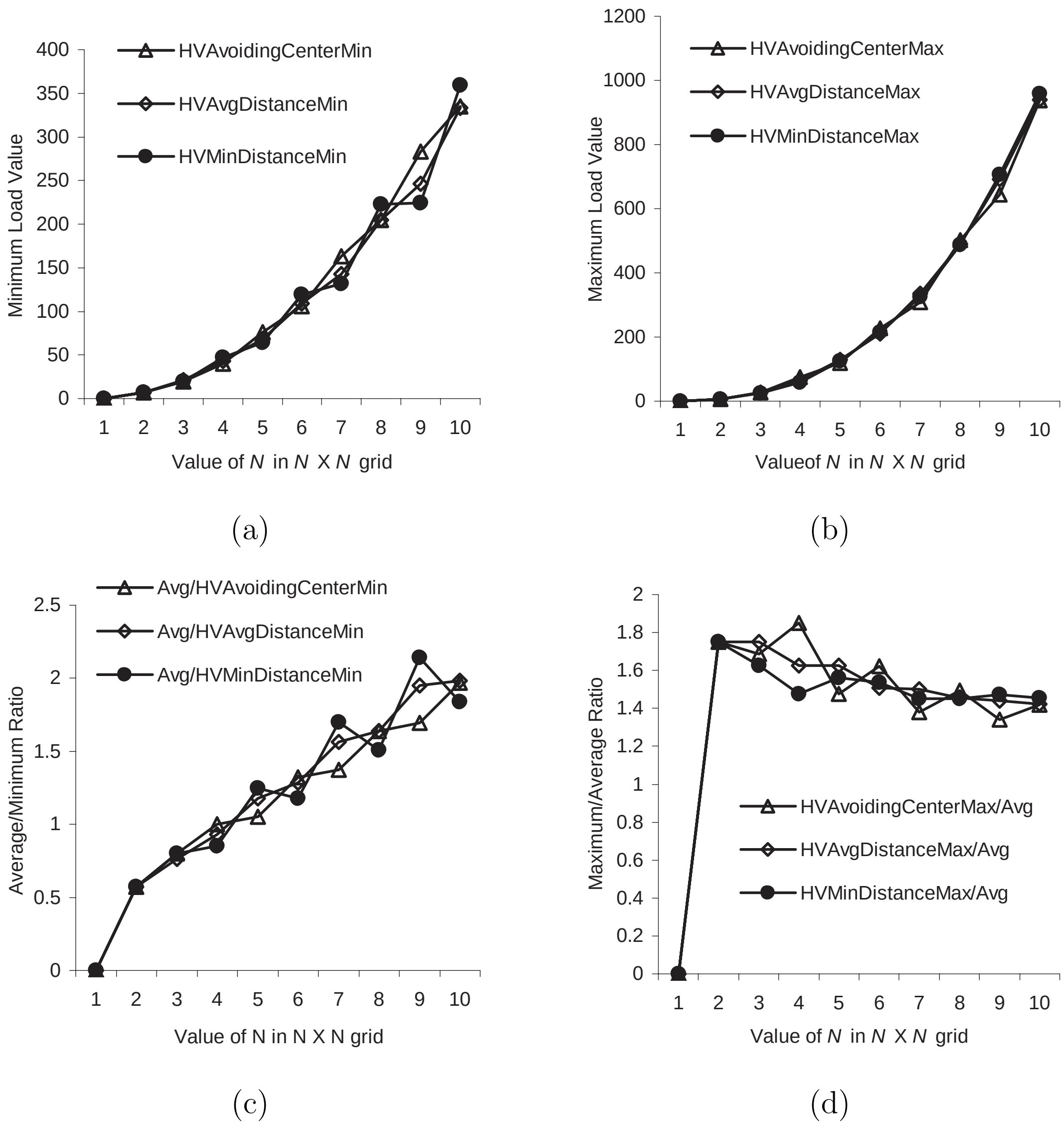


Fig. 21. Comparison of HVM-2,4,5 Variants (a) for Minimum Load, (b) for Maximum Load, (c) for Average/Minimum Load Ratio, (d) for Maximum/Average Load Ratio.

CHAPTER VI

ROUTING STRATEGIES IN DYNAMIC NETWORK SCENARIO

In sensor networks, nodes fail temporarily for a certain period of time. Hence, the grid network constantly has a dynamic configuration with nodes failing at different positions. This kind of situation is difficult to model. Therefore, we simulate this dynamism of the network by taking different configurations with nodes failed at different positions. We assume that once the grid configuration is decided and all to all communication phase begins, then no more nodes fail. We implemented *Zigzag*, *H-V*, *Servetto*, *Backtracking* and a *Hybrid-Method*(combination of *H-V* and *Servetto* methods) and compared them. Below is provided a brief description of the different methods we implemented. We considered *success-ratio* to be the criterion for deciding the efficiency of the methods.

VI.A Routing Strategies

Routing strategies studied in this scenario are described below.

1. *Horizontal-Vertical Method*: In this method, the messages are routed using *H-V* method as described in Section V. The only difference from static case is that this scenario contains failed nodes. At every node, the message may be routed through two nodes. If one of the nodes is functional, then the message will be routed through that node. When both the nodes fail, then the message does not reach its destination and is considered to be *failed* message. We consider the single source–single destination scenario and derive the probability of success for a message to be successfully received by the destination.

Lemma 8. *Denoting P(i,j,k,l) as the probability of success of path from node (i, j) to node (k, l) and p as the probability of failure of node in the grid, probability of success of H-V path in a single source–single destination mode is given from the recursive equations below:*

$$P(k, l, k, l) = 1 \tag{6.1}$$

$$P(x, l, k, l) = (1 - p)P(x + 1, l, k, l), 0 \leq x < (N - 1) \tag{6.2}$$

$$P(k, x, k, l) = (1 - p)P(k, x + 1, k, l), 0 \leq x < (N - 1) \tag{6.3}$$

$$P(i, j, k, l) = (1 - p)P(i, j + 1, k, l) + p(1 - p)P(i + 1, j, k, l) \tag{6.4}$$

Proof. Equation 6.2 is the probability of success of path from node to itself, which is always one. When the message is on the same column or on the same row as the destination, then it can no longer move in horizontal direction or vertical direction respectively. It is forced to move in vertical (if in same column) or horizontal (if in same row) direction. This movement is successful, provided the next node is functional, which is possible with a probability of $(1-p)$. From next node it follows the same scenario. Hence, we have Equations 6.3, 6.4. Now Equation 6.4 is derived as follows: In *H-V* path, we take horizontal movement preference over vertical movement and will move in vertical direction only if horizontal movement is not possible due to node failure. Message from node (i, j) has two nodes namely - either $(i, j + 1)$ or node $(i + 1, j)$ as next nodes. Hence the total probability of success of path from (i, j) to (k, l) is a weighted combination of probability of reaching the destination (k, l) from either of the two nodes $(i + 1, j)$ or $(i, j + 1)$. We have probability of reaching destination through node $(i, j+1)$ is given by (1-p) P(i,j+1,k,l), where (1-p) is attributed to

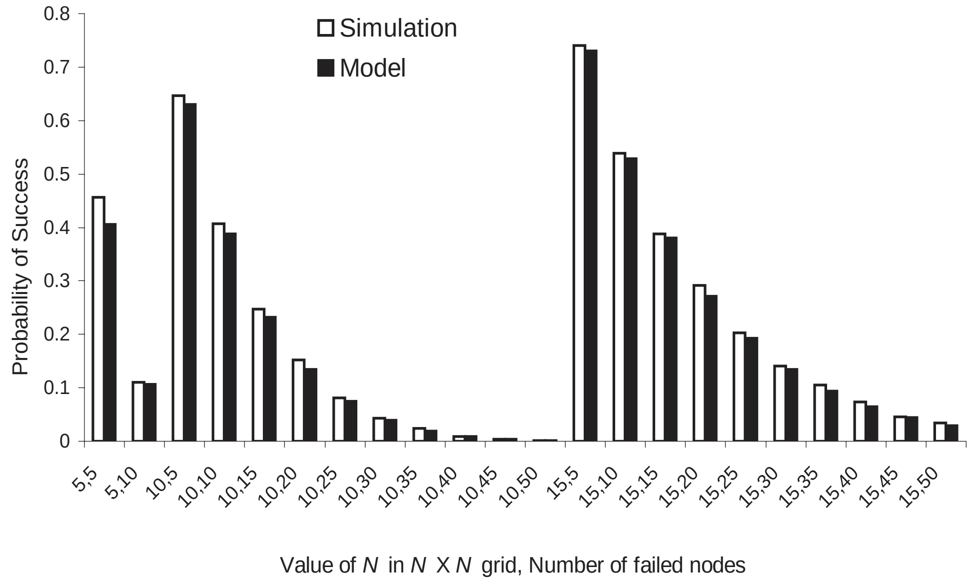


Fig. 22. Success-ratio Model in H-V Routing

the condition that $(i, j+1)$ should be functional. Due to horizontal preference, message is not routed through $(i+1, j)$ unless node $(i, j+1)$ happens to be failure. Hence, probability of reaching the destination through node $(i+1, j)$ is given by $p(1-p)P(i+1, j, k, l)$. By adding these two probabilities, we get the probability of success of path from node (i, j) to node (k, l) as given in Lemma 6.4. □

The above single source–single destination scenario is simulated and the model is confirmed as shown in Figure 22.

2. *Zigzag method*: Messages are routed using the *Zigzag* method in this scenario. When the message can no longer be routed from some node on its path to destination, then it is considered to be failed message. In [12], Badr *et al.* consider this strategy in their paper. In similar way to that derived in their papers, we derived the probability of success of a path in a single source–single

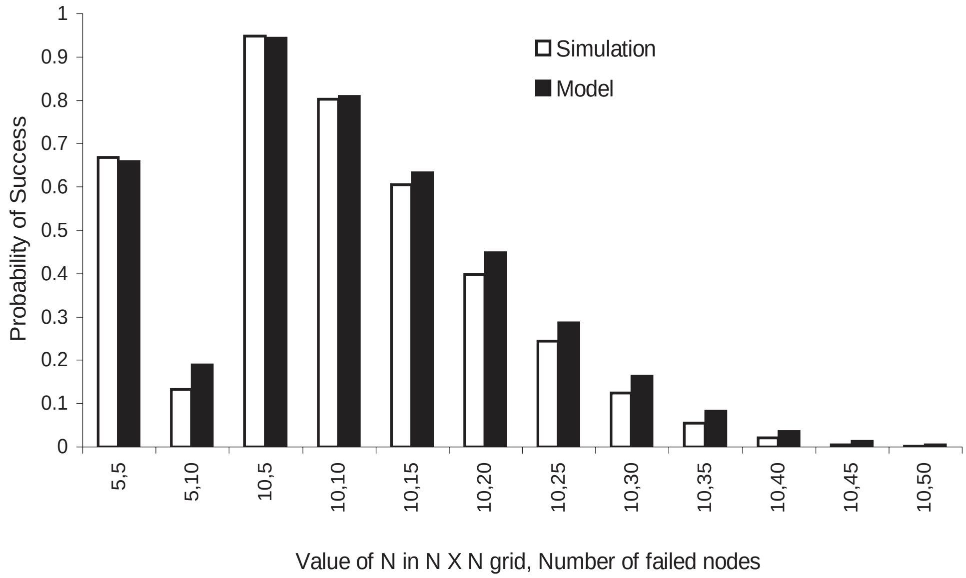


Fig. 23. Success-ratio Model in Zigzag Routing

destination mode using *Zigzag* routing.

Lemma 9. *Denoting P(i,j,k,l) as the probability of success of path from node (i, j) to node (k, l) and p as the probability of failure of node in the grid, probability of success of Zigzag path in a single source–single destination mode is given from the recursive equations below:*

$$
\begin{aligned}
P(k,l,k,l) &= 1 && (6.5)\\
P(x,l,k,l) &= (1-p)P(x+1,l,k,l), 0 \le x < (N-1) && (6.6)\\
P(k,x,k,l) &= (1-p)P(k,x+1,k,l), 0 \le x < (N-1) && (6.7)\\
P(i,j,k,l) &= (1-p)^2 P(i+1,j+1,k,l) + p(1-p)P(i,j+1,k,l) && (6.8)\\
&+ p(1-p)^2 P(i+1,j,k,l) + p(1-p)^2 P(i+2,j,k,l)
\end{aligned}
$$

Proof. Proof for *Zigzag* is similar to that shown for Lemma 8. We will derive it on similar lines. The path from $(i+1, j+1)$ to (k, l) follows again a *Zigzag* method if $(i+1, j+1)$ is reached from (i, j) through the node $(i, j+1, k, l)$. Hence the coefficient of $P(i+1, j+1, k, l)$ in the equation is a product of probabilities that both nodes $(i, j+1, k, l)$ and node $(i+1, j+1, k, l)$ are functional. Similar arguments lead to the coefficients of other nodes. $\square$

The model is confirmed using simulations as shown in Figure 23.

3. *Servetto Method*: In this method, the messages are routed using *Servetto* routing as described in [9]. In case they encounter the failed nodes, then the messages are not forwarded in that direction.

4. *Backtracking methods*: Backtracking methods have a special feature that when the process can no longer proceed in forward direction, then it can retract a step and proceed in a different direction towards the goal. We implemented backtracking method in both, *Horizontal-Vertical*, and *Zigzag* methods. The success-ratio obtained in these two methods is the same owing to the property, that they deliver the message successfully to destination whenever there exists a route from source to destination. They exhaustively walk through all the paths possible from source to destination, until they reach the destination.

5. *Hybrid Method*: This method is a combination of *Horizontal-Vertical*, and *Servetto* methods. In this method, the routing method is selected based on a parameter α which is selected prior to beginning of all to all communication phase. For every source destination pair, a random value is generated. If the random value is less than or equal to chosen alpha value, then *H-V* method is selected, else *Servetto* method is used as routing method between that source and destination. This method is described in [3] as described below:

$$\alpha HV method + (1-\alpha)\, Servetto \tag{6.9}$$

We implemented and simulated the values for different values of α (namely-0.1, 0.5, and 0.9) and for different values of failed nodes (namely-40, 100, and 200) in a 20×20 grid network.

VI.B Comparison of Routing Strategies

The different strategies mentioned in Section VI.A are simulated and their performance is compared considering *success-ratio* as the performance criterion. *Success-ratio* is defined as the ratio of successful messages transmitted from source to destination to the total messages (sum of *successful* messages and *failed* messages). In this calculation, we exempted from failed messages, all those messages which originated from or sent to failed nodes. These kind of messages are neither failed messages nor successful messages. The simulations are run for 400 times on a 20×20 grid network for each value of α with variable number of failed nodes. Then the average success ratio percentage of each method is calculated. α takes values - 0.1, 0.5, and 0.9. We made 40(10%), 100(25%), 200(50%) nodes of the network to fail and performed the simulations. Figures 24, 25, 26 show the results.

Figure 24 shows the results with 40 failed nodes in the network and for α taking values of 0.1, 0.5, and 0.9. The backtracking methods are not of great significance as they flood the network until the message reaches its destination. Hence, it is not considered a good strategy. Excluding backtracking method, *Zigzag* method does not flood the network and is better in performance than other methods. As shown in Figure 24, *Zigzag* is better than *H-V*, and *Servetto* methods by 3.5%, and 22.4% respectively. As the value of α is increased from 0.1 to 0.9, the performance of *Hybrid* method decreases. *Success-ratio* percentage of *Zigzag* exceeds that of *Hybrid-Method* by 5.7%,14.8%, and 23.8% with α taking values of 0.1, 0.5, and 0.9 respectively.

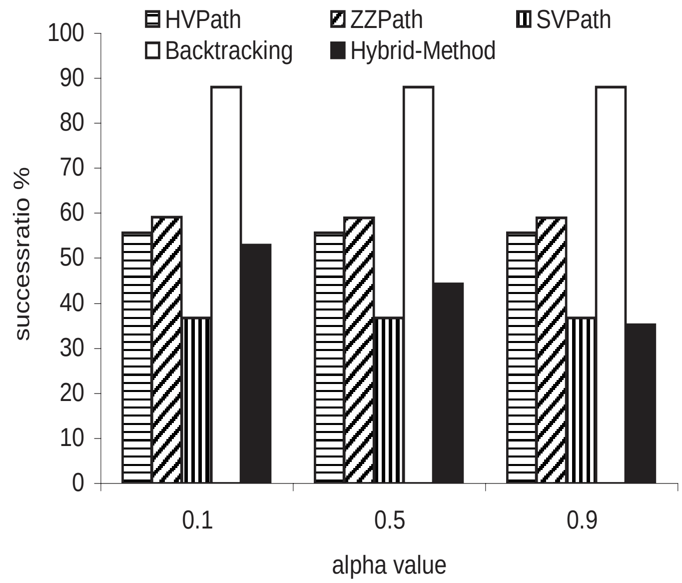


Fig. 24. Comparison of Success-ratio Percentages for Different Values of Alpha (p=10%)

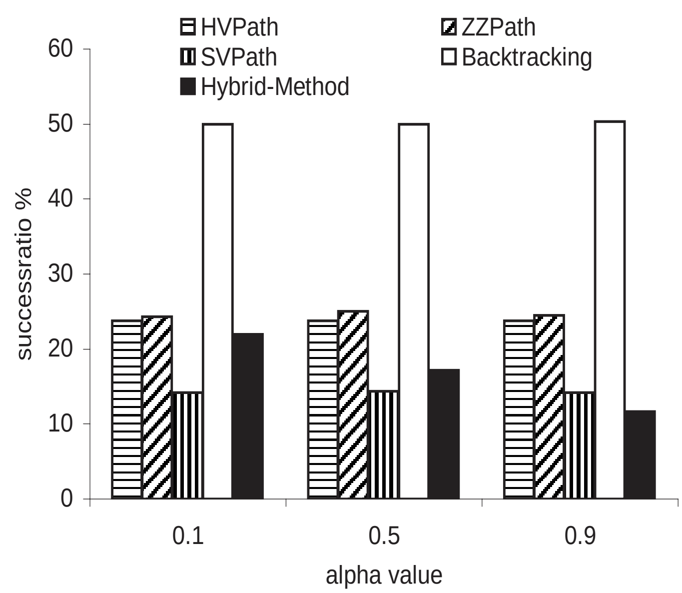


Fig. 25. Comparison of Success-ratio Percentages for Different Values of Alpha (p=25%)

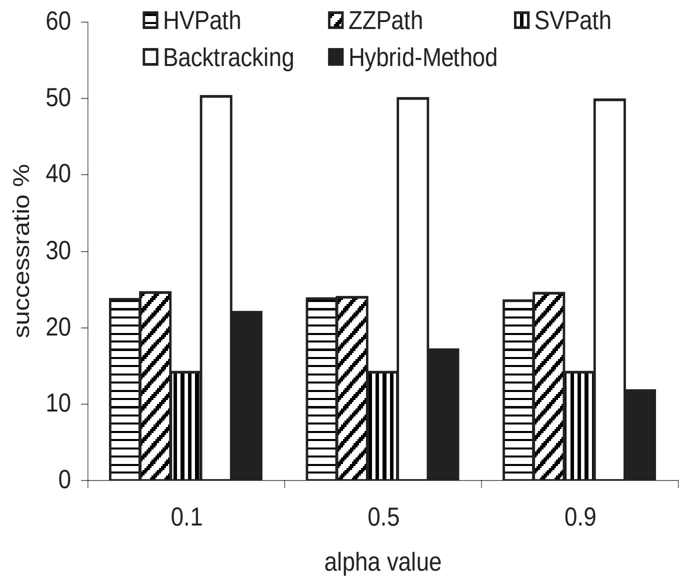


Fig. 26. Comparison of Success-ratio Percentages for Different Values of Alpha (p=50%)

Now we increased the number of failed nodes to 100(25%) nodes in the network. Results are shown in Figure 25. The *success-ratio* percentage of *Zigzag* is same as that of *H-V* method, and 10% better than that of *Servetto* method. Performance of *Zigzag* is better than that of *Hybrid-Method* by 2%,7%, and 13% for 0.1, 0.5, 0.9 values of α respectively. Now we increased the number of failed nodes to 200(50%) nodes in the grid and the results are shown in Figure 26. These results are similar to the results shown in Figure 25 with 100 failed nodes in the grid.

CHAPTER VII

CONCLUSION

In this thesis, we studied shortest path routing algorithms namely-*Horizontal-Vertical* and *Zigzag* routing. We analyzed *Servetto* method and showed that *H-V* method performs better load balancing of the grid network than *Servetto* strategy, in an *All to All* communication mode. We derived mathematical representations for the *maximum* and *minimum* loads on a static sensor grid, when these different routing strategies are employed in an *All to All* communication mode. We modified the *H-V* algorithm to obtain refined *H-V* strategies. In the dynamic network scenario, we studied the *Backtracking* strategy and a *Hybrid-Method* (combination of *H-V* and *Servetto*) along with the strategies mentioned in static network scenario. We derived formulae for the probability of success of path in a single source–single destination mode for specific routing strategies. Considering *success-ratio* as the performance criterion, the performance of *Zigzag* approach is better than that of *H-V*, and *Servetto* methods by 1%, and 14% respectively. It is better than *Hybrid* method by an average of 3.2%, 9.7%, and 16.6% for 0.1, 0.5, 0.9 values of α respectively. We showed through simulations that *Zigzag* routing performs better than other methods in a dynamic network with *success-ratio* as the performance criterion. We observed that the *Servetto* method proposed does not perform well in both the scenarios. In addition, the studied methods perform better than *Hybrid* method proposed recently.

Sensor Networks is an emerging field with a lot of potential for research. In this thesis we have dealt with deterministic routing strategies. Future work in this area may progress in the direction of designing adaptive strategies which would route messages based on the feedback from the network.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” *Computer Networks*, 38(4):393-422, March 2002.

[2] J. N. Al-Karaki, and A. E. Kamal, “A Taxonomy of Routing Techniques in Wireless Sensor Networks,” *Sensor Networks Handbook*, M. Ilyas and I. Mahgoub (eds.), Boca Raton, Florida, USA: CRC Press, 2004.

[3] G. Barrenechea, B. Beferull-Lozano, and M. Vetterli, “Lattice Sensor Networks: Capacity Limits, Optimal Routing and Robustness to Failures,” *Proceedings of the third international symposium on Information processing in sensor networks*, pp. 186–195, Berkeley, California, April 2004.

[4] C. L. Barett, S. J. Eidenbenz, L. Kroc, M. Marathe, and J. P. Smith, “Parametric Probabilistic Sensor Network Routing,” *Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, pp. 122–131, San Diego, California, September 2003.

[5] M. Kaufmann, J. F. Sibeyn, and T. Suel, “Derandomizing Algorithms for Routing and Sorting on Meshes,” *Proceedings of the 5th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 669–679, Arlington, Virginia, 1994.

[6] D. Braginsky, and D. Estrin, “Rumor Routing Algorithm for Sensor Networks,” *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pp.22–31, Atlanta, Georgia, 2002.

[7] C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks,” *Proceedings of the 6th*

annual international conference on Mobile computing and networking, pp. 56–67, Boston, Massachusetts, August 2000.

[8] B. Krishnamachari, D. Estrin, and S. Wicker, "Modelling Data-Centric Routing in Wireless Sensor Networks," *Proceedings of the 2002 IEEE INFOCOM*, New York, NY, June 2002.

[9] S. D. Servetto, and G. Barrenechea, "Constrained Random Walks on Random Graphs: Routing Algorithms for Large Scale Wireless Sensor Networks," *Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pp. 12–21, Atlanta, Georgia, September 2002.

[10] F. Ye, H. Luo, J. Cheng, S. Lu, and L. Zhang, "A Two-Tier Data Dissemination Model for Large-scale Wireless Sensor Networks," *Proceedings of the 8th annual international conference on Mobile computing and networking*, pp. 148–159, Atlanta, Georgia, September 2002.

[11] S. Dulman, T. Nieberg, J. Wu, and P. Havinga, "Trade-Off between Traffic Overhead and Reliability in Multipath Routing for Wireless Sensor Networks," *WCNC Workshop*, vol. 3, pp. 1918- 1922, New Orleans, March 2003.

[12] H. G. Badr, and S. Podar, "An Optimal Shortest-Path Routing Policy for Network Computers with Regular Mesh-Connected Topologies," *T-COMP(38)*, 1989.

[13] S. Shakkottai, R. Srikant, N. Shroff, "Unreliable Sensor Grids: Coverage, Connectivity, and Diameter," *IEEE INFOCOM - The Conference on Computer Communications*, vol. 22, no. 1, pp. 1073–1083, San Francisco, California, April 2003.

[14] C. M. Fiduccia, and P. J. Hedrick, "Edge Congestion of Shortest Path Systems

for All to All Communication," *IEEE Trans. on Parallel and Distributed Systems* 8(10): 1043-1054, 1997.

[15] J. Gao, and L. Zhang, "Load Balanced Short Path Routing in Wireless Networks," *IEEE INFOCOM 2004 - The Conference on Computer Communications*, vol. 23, no. 1, pp. 1099-1108, Hong Kong, March 2004.

[16] H. Dai, and R. Han, "A Node-Centric Load Balancing Algorithm for Wireless Sensor Networks," *GLOBECOM 2003 - IEEE Global Telecommunications Conference*, vol. 22, no. 1, pp. 548-552, San Fransico, California, Dec 2003.

[17] S. Lee and K. G. Shin, "Interleaved all-to-all reliable broadcast on meshes and hypercubes," *IEEE Transactions on Parallel and Distributed Systems*, vol. 5, no. 5, pp. 449–458, May 1994.

[18] R. Thakur and A. Choudhary, "All-to-All Communication on Meshes With Wormhole Routing," *Eighth International Parallel Processing Symposium*, pp. 561–565, Cancun, Mexico, April 1994.

[19] S. Hinrichs, C. Kosak, D. R. O'Hallaron, T. M. Stricker and R. Take, "An Architecture for Optimal All-to-All Personalized Communication," *Proceedings of the sixth annual ACM Symposium on Parallel Algorithms and Architectures*, pp. 310–319, Cape May, New Jersey, 1994.

[20] Y. Yang, and J. Wang, "Pipelined All-to-All Braodcast in All-Port Meshes and Tori," *IEEE Transactions on Computers*, vol. 50, no. 10, pp.1020–1032, October 2001.

[21] D. S. Scott, "Efficient All to All Communication Patterns in Hypercube and Mesh Topologies," *Proc. of* 6^{th} *IEEE Distributed Memory Computing Conference*, pp.

398–403, Portland, Oregon, April 1991.

VITA

Lakshmana Prasanth Nittala Venkata has done his B.Tech in computer science and engineering from Andhra University, Visakhapatnam, Andhra Pradesh, India. He joined the Department of Computer Science, Texas A&M University as a masters student in September, 2002. His Permanent Address is:

S/o Sri. N. V. Ratnam,
Door No 1-31, Baji Junction,
Visakhapatnam, A.P.,
India-530027.

The typist for this thesis was Lakshmana Prasanth Nittala Venkata.